Elucidating self-assembly and antimicrobial strategies of synthetic peptides an in silico investigation

Marzuoli, Irene

Awarding institution:
King's College London

Elucidating self-assembly and antimicrobial strategies of synthetic peptides: an in silico investigation

Irene Marzuoli

Randall Centre for Cell and molecular Biophysics

King's College London

This dissertation is submitted for
the degree of Doctor of Philosophy

October 2019

To my parents, for letting me go

Declaration

This dissertation describes work I have carried out between October 2016 and September 2019 at the Randall Centre for Cell and Molecular Biophysics of King's College London, under the supervision of Prof. Franca Fraternali (first supervisor), Dr. Chris D. Lorenz (second supervisor, Department of Physics) and Dr. Maxim G. Ryadnov (industrial supervisor, National Physical Laboratory).

The work was supported by the Engineering and Physical Sciences Research Council (EPSRC) through the Centre for Doctoral Training Cross Disciplinary Approaches to Non- Equilibrium Systems (CANES) (Grant No. EP/L015854/1).

This dissertation contains material appearing in the following articles:

- Kepiro, I. E., Marzuoli, I., Hammond, K., et al. (submitted and in second revision). Engineering chirally blind protein pseudo-capsids into nanoprecise antibacterial persisters.
  Chapter 3 contains results presented in this paper. Appendix A.1 contains the submitted version.

- Marzuoli, I., Margreitter, C., Fraternali, F. (2019). Lipid Head Group Parameterization for GROMOS 54A8: A Consistent Approach with Protein Force Field Description. *Journal of Chemical Theory and Computation*, 15(10):5175–5193
  This paper is presented in Chapter 4.

In addition to the above, I have contributed to the following publication during the course of my PhD:

- Milano, G., Marzuoli, I., Lorenz, C. D., and Fraternali, F. (2017). Self-assembly at the multi-scale level: challenges and new avenues for inspired synthetic biology modelling. In Ryadnov, M., Brunsveld, L., and Suga, H., editors, *Synthetic Biology: Volume 2*, pages 35-64. The Royal Society of Chemistry, London.

This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements. It has not been submitted in whole or in part for any degree or diploma at this or any other university.

Irene Marzuoli
October 2019

Acknowledgements

Many people contributed to my scientific and personal growth throughout my PhD, and this thesis would not have been possible without them.

First of all, Franca, an excellent scientist, supervisor and woman: she showed me how true enthusiasm for research and life can withstand thousands everyday difficulties.

Then, all the present and past members of the Fraternali's group: Anna, Carlos, Christian, Emma, Jamie, Joseph, Marius and Sun, each of you made the office a more enjoyable and scientifically stimulating place. A special thanks to Carlos and Christian, for the great (and sometimes unexpected) collaborations; and to the people who shared these years at the same pace: Joseph, whose cheerfulness made working days and scientific discussions more pleasant (not to mention the treats he supplied), and Marius, for the curious and caring attitude which granted rewarding, serendipitous, conversations (and lunches in the sun).

I am also grateful to my second supervisor, Chris, and all his group, for the concrete and prompt support on practical - or less practical - problems, whenever I needed it.

Thanks to Prof. Giuseppe Milano for hosting me and sharing his knowledge, and to Dr. Jens Kleinjung for help and suggestions in the initial stages of my project.

To all the members of the CANES community, students and academics, thank you for pushing me to stay updated on all aspects of science. In particular, I want to single out the "italian subset", for the warmth, the support, and the encounters at all the corners of the world. Especially, thanks to Claudio, for his presence in good and difficult moments, and all the jazz music; to Aldo, for challenging (and supporting) my ideas about research, universe and everything; and mostly to Carla, a colleague, a flatmate, a friend: navigating our job and every day life together made me a more serene person.

I owe my perseverance also to Laura, Celeste and Nicola, who, comparing their PhD experience with mine, formed an international, and interdisciplinary, support network.

Finally, my fondest gratitude goes to my parents. For encouraging and trusting me in this adventure, without asking back anything, a part from brief stories from my research life - those were the only presentations I truly enjoyed giving.

Abstract

The modern insurgence of antimicrobial resistance prompted the research of new drug alternatives. In parallel, the problem of their delivery has stimulated the research of novel biomimietic vehicles. Synthetic materials can be designed to perform both functions effectively. Recently engineered nanocapsules were shown to promote bacterial membrane poration and gene delivery into mammalian cells. Their constitutive molecule, capzip, is a three branched peptide, which contains sequences inspired from a naturally occurring antimicrobial peptide (AMP). AMPs act on bacteria disrupting their membrane, a mechanism which does not strongly promote resistance.

As the atomistic details of the capzip nanocapsule assembly are still unknown, this project studied its structure in water and its interaction with model membranes, by means of multiscale Molecular Dynamics simulations. The *in silico* investigation clarified the preferred structure that capzip adopts in order to form robust capsules. In particular, the original formulation of capzip included an amphiphilic pattern to promote antimicrobial activity, but simulations proved that this has a key role in granting structural stability as well. This provides insight for the development of the next generation of multi-branched antimicrobial molecules. The multiscale investigation performed prompted also a comparison between coarse-grained force fields, which will contribute to inform the choice of the most adapt one for future simulations of large peptidic assemblies.

The structures found to be stable in solution were selected for further simulations at the interface with a model bacterial and mammalian membrane. The insertion of charged residues in the membrane ester region produced a local decrease in lipid mobility and, under the effect of an externally applied electric field in the physiological range, pore formation with subsequent membrane disruption. Coarse-grained simulations confirmed these findings, clarifying the attraction mechanism between the capsule and the model bacterial membrane. Moreover, they suggested a lower affinity with the model mammalian membrane. The exploration of the peptide-membrane interactions prompted an investigation of the currently used lipid parameters in the GROMOS atomistic force field. Given the inconsistency between the parametrisation of proteins and lipids found in the latest versions of the force field, we proposed a new parametrisation that reconcile these. The new parameters showed a peptide-membrane interaction which is less biased by the simulation's initial conditions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

SCIENCE has always thrived thanks to human curiosity and to the desire of modifying the environment around us. Out of the many science fields at service of human wealth, the most challenging and still far away from being mastered is the comprehension and manipulation of the human body and mind. It is striking how we are finally scratching the understanding of these two entities in the same historical moment, at a point where computers start imitating human reasoning [Editor BBC, 2016; Google, 2016; Alom et al., 2019] and biological materials are turned into semi-functional organs [Rossi et al., 2018].

The process of understanding is intertwined with the process of engineering, and goes in parallel with the development of technology. As such, in the past two centuries we witnessed a technological evolution that condensed the collective expertise we accumulated so far into objects belonging, by now, to our daily life.

This boost of technical knowledge allowed for the development of novel instruments, and in turn for the accumulation of a broader scientific knowledge. Consequently, the times of all-round scientists left the way to an era in which specialisation is necessary to master and understand a subject, in the hope and trust that piecewise knowledge builds an organic body once the effort of many scientists is joined together.

The increasing understanding of how the body works - or fail to do so sometimes - together with the discovery of bacteria and the symbiotic or disruptive relationship we have with them, poses novel questions to science: why the body, which is carefully design to function effectively, is susceptible to pathogenic diseases, malfunctioning of its organs or failures in controlling its cells growth (tumours)? And why in some cases

it defends itself also against beneficial drugs? But mostly, what can we do to repair such faults?

Our biological and emotional push to change the course of nature to improve our lives lead to the development of "medical technologies" such as drugs and treatments to aid the body in fighting the agents which attack it. Historically, this proceeded from a process of trial and error of natural substances, to an informed synthesis of artificial drugs to repair the body [Wishart et al., 2018], together with disinfectants [City of Vienna Climate Protection Programme, 2012] and antibiotics [The Johns Hopkins University, 2017] to fight pathogens: all together, we are closer and closer to the *magic bullet* envisioned a century ago by Nobel Prize Paul Ehrlich, who dreamed of a "personalised and tailored drug" able to target specific molecular defects, while being beneficial to other healthy cells [Strebhardt and Ullrich, 2008]. A success of the kind will be a life saving technology, and condense in a single "bullet" centuries of efforts to understand nature.

Given the vast amount of knowledge on the topic, every research task focusses on single, simplified questions to complete the full picture. While new problems arise and are answered, new experimental techniques are developed to investigate them. However, the difficulties of studying micrometric systems as cells and bacteria are manifold, as we often do not possess instruments to look at them with the desired level of detail, or to do so without perturbing their natural conditions. Therefore, in the last decades a new investigation approach emerged, proposing to model the systems of interest from the theoretical knowledge gathered so far, in terms of their structure, behaviour and properties. Computational Biology is the field that proposes to do that, implementing those models first and then querying them for properties which are still unknown. At times, this proves to be the only method possible to answer the question posed, for example to uncover the atomistic details of protein unfolding, to monitor the dynamics of transport channels or instead to predict the evolution of a cancer cell population [Lee et al., 2009; Dror et al., 2012; Wang et al., 2015b].

This thesis aims at employing techniques belonging to Computational Biology to answer the specific question of how an artificial molecule behaves within selected biological environments. Thus, while studying a structural biology problem, it places

itself at the boundary with the fields of Bioengineering and Medicine: the following introduction is meant to give an overview of the many different challenges these fields have faced in recent years and the solutions found to those challenges, motivating the interest in the development of novel molecules as the one which will be described in this work. In particular, two problems are highlighted: a) the insurgence of drug and antimicrobial resistance, and b) the problem of drug delivery, as they are both addressed by the system in exam. Figure 1.1 provides a work flow of this introductory chapter, to help the reader in identifying the sections of interest.

1.1 Antimicrobial resistance

For most of the last century, the development of new drugs rotated around the paradigm that a drug is a small inorganic compound (of mass up to 900 Da), which intervenes on a specific target of a mammal or bacterial cell. Very often the targets of interest are (intracellular) proteins: out of the 695 small drugs approved by FDA (the American Food and Drug Administration agency) to target human molecules, 667 acts on proteins. Similarly, 189 of the 198 approved to treat pathogens have a protein as their target (with all the caveats coming from the challenges of identifying an unambiguous target, especially when the drug binds to a protein complex or to a number of closely related gene products [Santos et al., 2017]).

In presenting the aforementioned figures, the data were naturally split among the drugs which act on human molecules, "repairing" a faulty process in the human body, or the ones active against bacteria (commonly named antibiotics) which "disrupt" the bacterium life cycle in order to kill or prevent the reproduction of the pathogen. It appears evident that the pool of drugs available to the second purpose has a substantially lower number of compounds than the ones addressing human molecules. This comes from the nature of the action they perform: molecules targeting human proteins need to be highly specific to avoid interference with other proteins or with healthy cells, and in a sufficient number to address the variety of diseases affecting the human body. Antibiotic must be non-toxic for human cells as well, i.e. their target must not be shared between mammal and bacterial cells, but there is a less stringent requirement

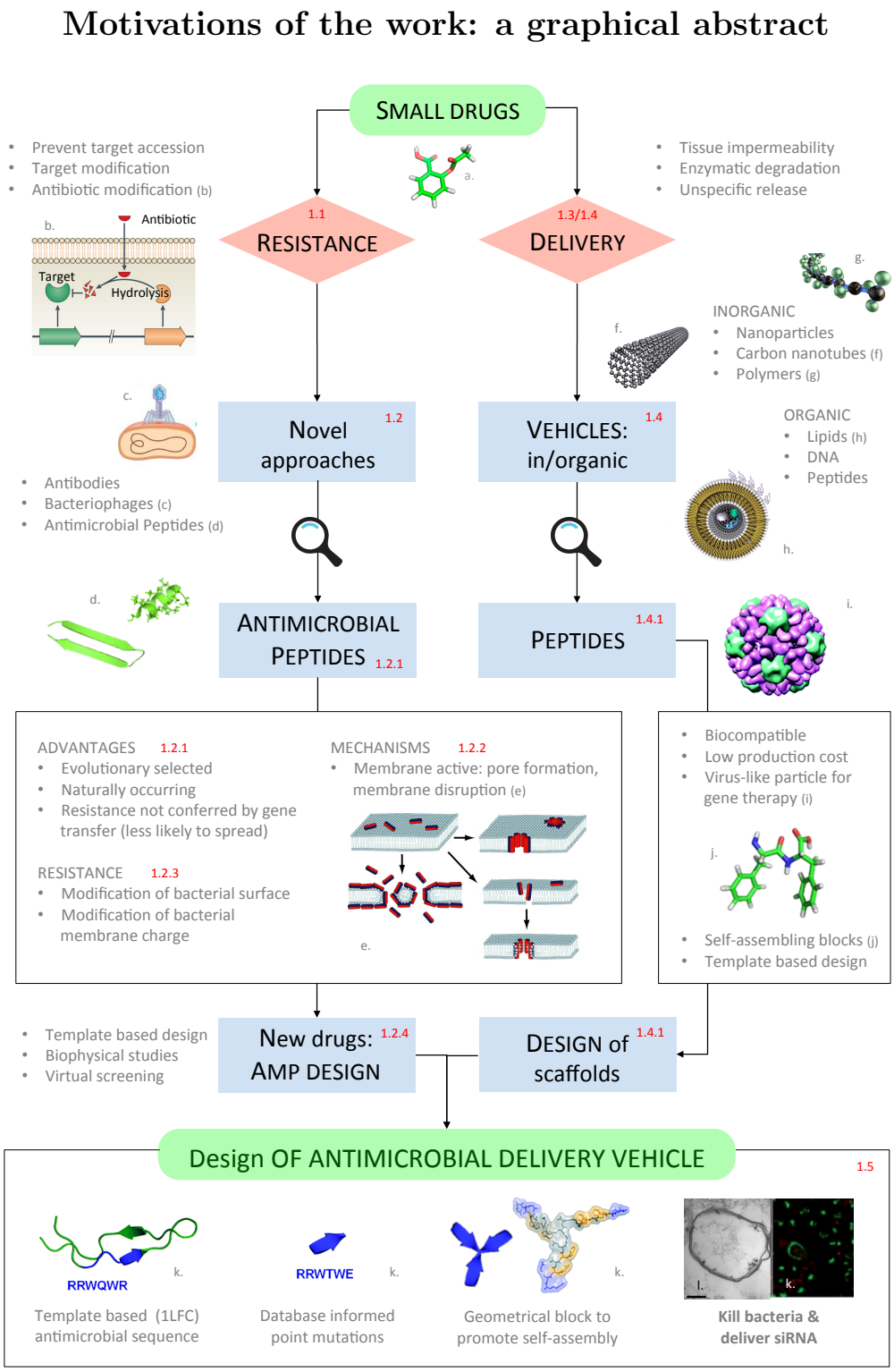


Figure 1.1: Figures a. Acetylsalicylic acid and j. Diphenyl-alanine in bond representation [VMD software Humphrey et al. [1996]]. Remaining figures adapted from: b. Blair et al. [2015]; c. Kim [2017]; d. Torres et al. [2019]; e.Nguyen et al. [2011]; f. The Innovation Society [2019]; g. Mettler Toledo [2018]; h. [Wikipedia, 2015]; i. Schoonen and van Hest [2014]; k. Castelletto et al. [2016]; l. Kepiro et al. [2019]. The numbers in red refer to the sections of this chapter in which the topic is treated in more detail.

on their selectivity against different bacterial species. On the contrary, it is often useful to have a broad-spectrum compound. This cross-species efficacy and at the same time non-toxic property is obtained thanks to the evolutionary relationship among bacterial species, and between bacteria and humans. While the first are closely related, and therefore share homologous proteins with very similar structures, humans have less folds in common with them, allowing for a resilience against bacteria-targeting drugs. The full picture is actually more complex, as the set of bacterial species is very diverse, and the cross-species effectiveness of some drugs does not extend to the whole bacterial populations. This can be a positive feature, given the large amount of beneficial bacteria that live in symbiosis with the human body (especially in the gut) and that must be preserved for an optimal wellness.

Nevertheless, in the framework described above, it is understandable that first-time research on antibiotics was satisfied with the development of a handful of potent, broad-spectrum compounds. Penicillin, the first to be synthetically produced, was isolated from a mould in 1928 by Alexander Fleming. It acts inhibiting the formation of a cross-links between particular molecules (peptidoglycan) in the bacterial cell wall, binding to the enzyme responsible for their catalysis, and thus preventing the complete wall formation [Gordon et al., 2000] (for further details on the bacterial cell membrane see Section 1.2.1). As foreseen from Fleming himself in his Nobel Prize acceptance speech, some species of bacteria quickly became immune to penicillin, and this was achieved in many ways: either by production of an enzyme that degrades penicillin, by subtle changes in the structure of the penicillin-binding proteins to prevent such binding, or again by removal of the drug from the cell through specially re-purposed efflux pumps [Lobanovska and Pilla, 2017].

The mechanisms just outlined are not an exceptional characteristic of penicillin, and many drugs lost their effectiveness against some bacteria from their discovery, urging the research of new ones on a constant basis. By now, a broad knowledge has been gathered on how bacteria escape the action of a drug: this understanding helps interpreting the pitfalls of existing drugs and identifying the characteristics sought for new compounds.

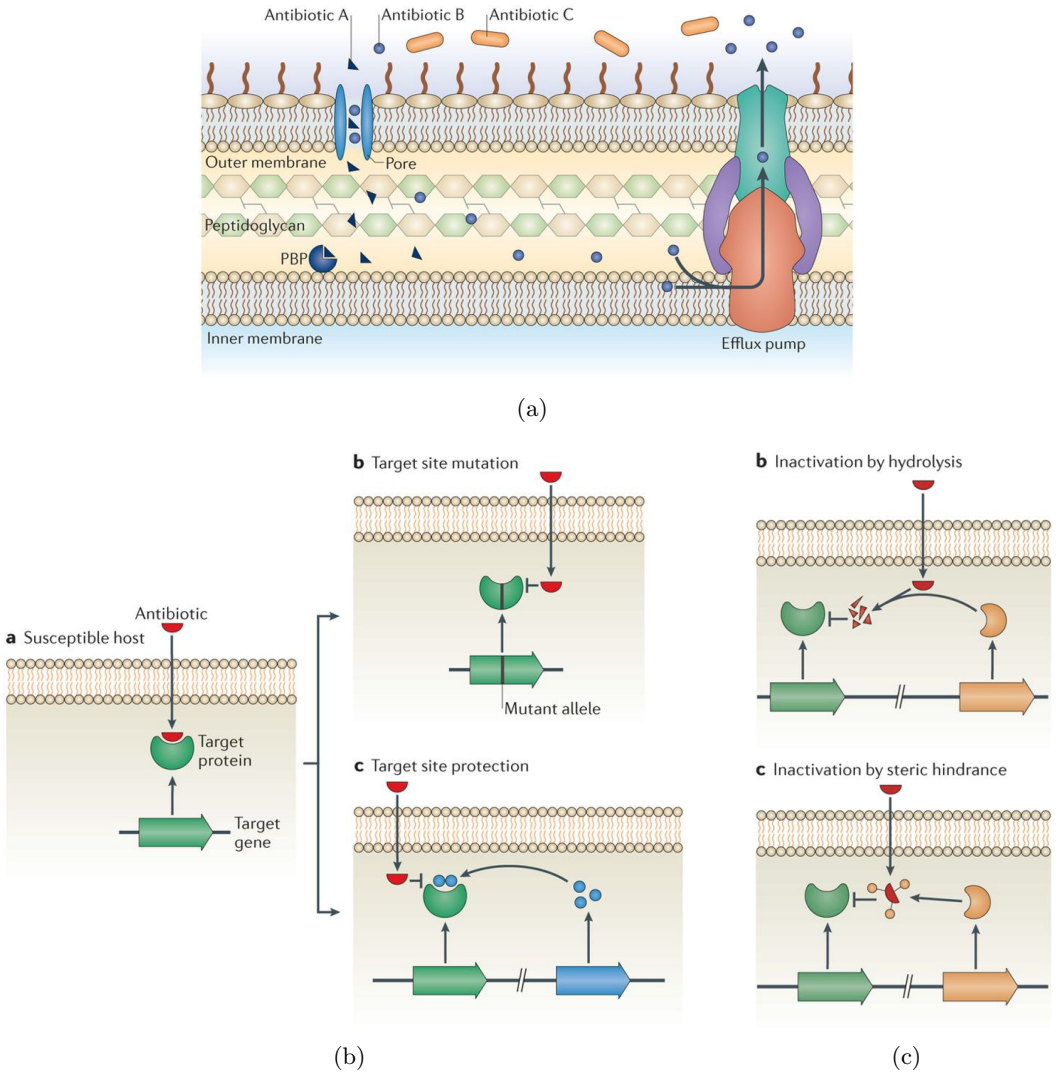


Figure 1.2: Mechanisms of antimicrobial resistance to small drugs. (a) Removal of antibiotic B by efflux pump and inaccessibility of antibiotic C to the Penicillin Binding Protein target because of membrane impermeability. (b) Target site change via mutation or protection. (c) Direct interactions with antibiotics causing its disruption or structural modification. Reproduced from Blair et al. [2015].

1.1.1 Mechanisms of antimicrobial resistance to small drugs

Antimicrobial resistance can manifest through many different mechanisms, which can be grouped in three main classes, in line with the three processes mentioned in the example of the penicillin resistant bacteria.

Prevention of access to target A first class of resistance mechanisms aims at minimising the intracellular concentration of the antibiotic preventing its penetration or maximising its efflux in the eventuality it has entered the cell (Figure 1.2(a)). Not all the

molecules can enter the cell permeating the membrane, and this holds particularly for hydrophilic antibiotics tackling Gram-negative bacteria, which are intrinsically poorly permeable because of the presence of a double membrane [Delcour, 2009] (see Section 1.2.1). These molecules must then be imported into the cell through outer-membrane porin proteins [Vargiu and Nikaido, 2012; Kojima and Nikaido, 2013]. Resistance arise when porins are either replaced with more selective channels, which prevent the antibiotic penetration, or down regulated so that the internal concentration of the drug does not reach a critical concentration [Lavigne et al., 2013]. Porin-coding genes can also accumulate multiple mutations, to acquire the selectivity they lack in their wild type [Poulou et al., 2013].

A complementary strategy to prevent drug influx is to employ bacterial efflux pumps. Some of them are denominated multidrug resistance (MDR) efflux pumps for their effectiveness in the task and are produced by many bacteria [Floyd et al., 2010; Ogawa et al., 2012]. Over-expression of such efflux pump is observed in multidrug-resistant bacteria, triggered by exposition to the drug. Such upregulation proceeds via mutation in the relative regulatory network [Abouzeed et al., 2008], or simply as a response to environmental signals [Nikaido et al., 2011]. Additionally, the genes coding for them can be transferred via plasmids to other bacterial species [Dolejska et al., 2013]. Indeed, bacteria are able to exchange genetic material with other individuals via small rings of DNA in a process called conjugation [Sørensen et al., 2005], so that advantageous resistant genotypes can spread quickly across species.

Change or modification of the antibiotic target The second class of resistance mechanisms works modifying the antibiotic target: most antibiotics bind to their substrate with high affinity and specificity, thus small modifications in the target structure can disrupt an efficient binding, still allowing the target to maintain its normal function (Figure 1.2(b)).

Mutations of some residues in the binding pocket (upon mutation in the gene coding for it) or its post-translational protection via addition of chemical groups are equally wide spread strategies. Notable examples of the first include the development of methicillin- (an antibiotic in the penicillin class) resistant strains of *S. aureus* [Shore

et al., 2011; Billal et al., 2011]. Again, several of these mutations are acquired by horizontal gene transfer from other bacterial species. For post-translational protection, the most relevant mechanism of chemical group addition is methylation, which, for example, is very common when the drug target are rRNA subunits [Long et al., 2006]).

Direct modification of antibiotics Finally, bacteria can destroy drugs, usually by hydrolysis, or modify them by transfer of a chemical group (Figure 1.2(c)). The first drug-degrading enzyme discovered was penicillinase [Abraham and Chain, 1988; Lobanovska and Pilla, 2017]. Since then, thousands of similar enzymes have been identified that can modify antibiotics of different classes [Livermore, 2008; Nordmann et al., 2011]: these enzymes co-evolves with newly developed drugs, to include in their spectrum of action new compounds of composition similar to the ones they were originally effective on [Woodford and Johnson, 2013].

Antibiotics constituted by large molecules with many exposed hydroxyl and amide groups are instead particularly susceptible to addition of chemical groups. Many enzymes are responsible for this, and according to the chemical moiety added they are grouped in acetyltransferases, phosphotransferases and nucleotidyltransferases [Wright, 2005].

1.1.2 Outlook on antimicrobial resistance

All together, the recent progress in understanding the mechanisms of antimicrobial resistance has helped in directing the development of new drugs. In particular, it has promoted the modification of existing compounds to escape the resistance developed by bacteria.

It must be noticed that many of the available drugs are bacteriostatic agents as opposed to bactericidal: i.e. they prevent the bacterium growth rather than kill it, as they are meant to slow down the damage while host defence mechanisms eradicate it. Thus, if an high dosage of a bactericidal agent may extinguish the bacterial population and eradicate the disease, bacteriostatic drugs allow bacteria to start again the reproduction cycle once removed (if the host defence could not properly work), and are thus more prone to “train” resistant bacteria.

The severity of the AMR threat is such that it has been raised to the status of national emergency in several countries, including UK. Strict regulations on the health, agricultural and food industry sector must be taken to prevent the misuse of antibiotics, as we are leaving the century in which antibiotics were discovered, to enter a phase in which we count the number of the ones loosing efficacy [O'neill, 2016].

1.2 Alternative antibiotic strategies: antimicrobial peptides

In the landscape sketched above, the development of novel drugs is of crucial importance. The progress in understanding the mechanisms of antimicrobial resistance has helped this process mainly promoting modification of existing compounds to escape the resistance developed by bacteria.

However, it would be even more beneficial to have at disposal a new paradigm for their design, in order to attack pathogens in a completely novel way, avoiding to target pathways which are known to lead easily to resistance. Several novel materials have been developed for the task, not to rely on small molecules and to exploit different mechanisms of action, for example antibodies, bacteriophages and antimicrobial peptides [Mantravadi et al., 2019].

The use of pathogen-specific antibodies relies on mechanisms of the host immune system, and bacteriophages therapy employs viruses which infect bacteria and archea rather than eukarya. Finally, some peptides can have an active role against bacteria, when their sequence possesses specific characteristics, and are thus referred to as antimicrobial peptides and they are the main focus of this thesis. The following subsections will explore their characteristics, modes of action, and the response of bacteria against them. It is indeed crucial to understand the knowledge available on these molecules versus the questions that are still open, in order to direct the efforts of future research. This holds in particular when the investigation proceeds by use of simplified models, as the ones employed in Computational Biology. Meaningful results can be obtained only if such modelling is performed in a sensible and informed fashion.

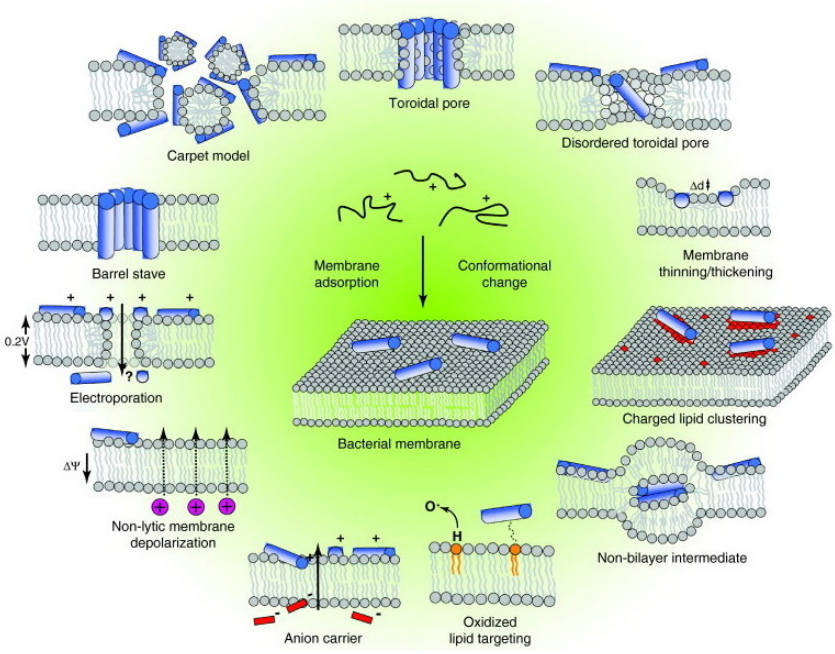


Figure 1.3: Events occurring at the bacterial cytoplasmic membrane following initial antimicrobial peptide (AMP) adsorption. Reproduced from Nguyen et al. [2011].

1.2.1 Membrane active peptides

Antimicrobial peptides (AMPs) are naturally produced by eukarya, either as stand-alone sequences or embedded in larger proteins, as a first, weak, and broad-spectrum defence against bacteria [Nguyen et al., 2011]. This pool of molecules has been selected through evolution to be active against pathogens, suggesting that they will most likely not cause resistance in a near future.

To exploit their potential and engineer AMP-like molecules, a careful characterisation and classification of such peptides must be done. This task has been carried on throughout the past decades, but because of its complexity at present there are many peptides with ascertained antimicrobial activity for which the mode of action is still not fully understood [Ebbensgaard et al., 2015]. However, some general characteristics of these sequences and some of the mechanisms they employ have emerged. Unsurprisingly, AMPs are heterogeneous in sequence, structure, targets and modes of action, to tackle the different challenges bacteria pose. Their size can vary between 6 and 59 amino acids [Brogden, 2005]: despite being small with respect to the average size of a protein in the human body, these macromolecules are hundreds of times larger than

small molecule drugs and as such they penetrate and act on bacteria differently with respect to them.

The most common target of AMPs is the bacterial membrane. Many of them cause disruption of the microbial membrane while others translocate into the cytoplasm to act on intracellular targets, and the combination of the two is not uncommon either [Hancock and Sahl, 2006] (Figure 1.3). In general, it is widely accepted that membrane interaction and its relative consequences are a key factor for the antimicrobial activity of AMPs [Nguyen et al., 2011].

As such, we propose a brief overview of the structure of the bacterial membrane [Silhavy et al., 2010], and of its differences with the one of mammalian cells, to better understand how AMPs can be effective and selective on bacteria at once.

Structure of bacterial membrane The determinant driving the interaction between AMPs and bacterial membranes is the positive charge that many AMPs present, opposed to the negative charge of the latter [Nguyen et al., 2011; Mahlapuu et al., 2016]. It is striking that such simple mechanism, based on the presence of a certain number of negatively charged lipids, holds across many bacterial species despite the great variability found in their membrane composition. Indeed, based on the differences in their cell envelope, bacteria are classified into two macro families, Gram-positive and Gram-negative. In Gram-positive bacteria, the cytoplasmic membrane is surrounded by a thick peptidoglycan layer, while for Gram-negative bacteria this membrane (which assumes the name of internal one) is surrounded by a thin peptidoglycan layer and an outer membrane [Silhavy et al., 2010; Lin and Weibel, 2016] (Figure 1.4).

Starting from the inside and proceeding outwards (from bottom to top of Figure 1.4), the cytoplasmic membrane of both Gram-positive and Gram-negative bacteria is rich in phospholipids like phosphatidylethanolamine, which is neutral, and phosphatidylglycerol, cardiolipin, or phosphatidylserine, which have negatively charged headgroups, highly attractive for positively charged AMPs [Silhavy et al., 2010; Lin and Weibel, 2016]. Simplified models of this membrane often opt for a 1:3 ratio between negative and neutral lipids, giving a value of roughly 0.5 e/nm^2. This is often sufficient to promote the preferential interaction between this membrane and the pep-

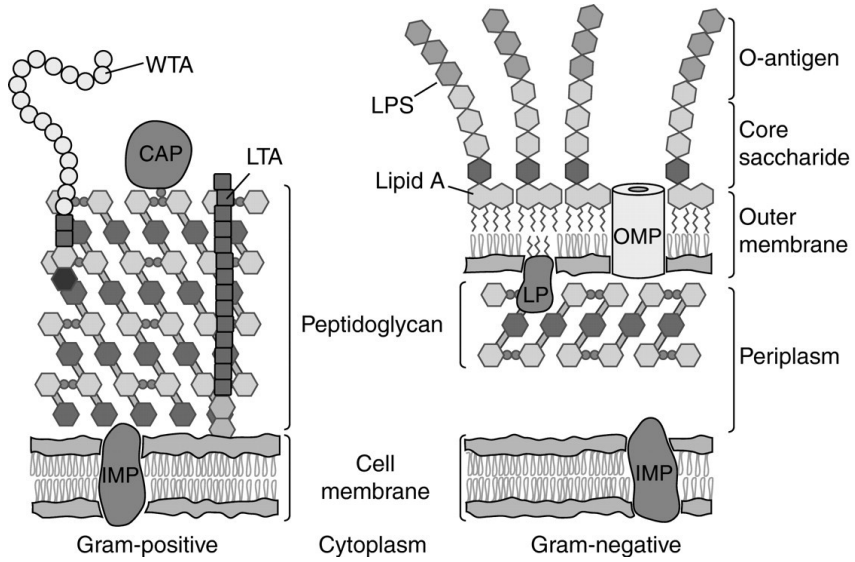


Figure 1.4: Structure of Gram-positive and -negative cell envelope. IMP: integral membrane protein; CAP: covalently attached protein; LTA: lipoteichoic acid; WTA: wall teichoic acid; LP: lipoprotein; OMP: outer membrane protein; LPS: lipopolysaccharide. Reproduced from Silhavy et al. [2010].

tides - provided they get into its proximity. Perturbation of this membrane is highly disruptive for the bacterium as many functions are associated to it: as bacteria do not possess organelles, all the membrane-related proteins reside and perform their function within the inner membrane.

In the case of Gram-negative bacteria (Figure 1.4, right), the inner membrane, together with the outer one, delimits the periplasm space, an aqueous cellular compartment, which allows the sequestration of harmful substances and the transport of nutrients. Inside the periplasm is situated the peptidoglycan cell wall. This substance is made of a disaccharide cross-linked by penta-peptide side chains, and these repeated units constitute the rigid skeleton of Gram-negative cells [Gan et al., 2008]. Peptidoglycan is fundamental for cell life as its damage results usually in living but not viable cells [Joseleau-Petit et al., 2007].

The outer membrane is covalently bonded to the cell wall through Brauns lipoproteins (or LP) [Asmar and Collet, 2018]. This membrane presents an asymmetric structure: phospholipids are present in the inner leaflet, while the outer leaflet is composed of glycolipids, mainly lipopolysaccharides (LPS) [Silhavy et al., 2010]. These complex molecules consist of lipid A, which presents multiple fatty acids, and a polysaccharide

[Raetz and Whitfield, 2002]. The polysaccharide is made of an inner core, covalently bonded to the lipid, an outer core, and finally a repetitive glycan polymer (O-antigen). The O-antigen (top right in Figure 1.4) is the molecule exposed by Gram-negative bacteria to the external environment and thus is the target of antibody recognition.

Given the complexity of the Gram-negative cell envelope, and especially the presence of the LPS layer, these bacteria are particularly impermeable to hydrophilic molecules, which are usually imported within the cell through porins and similar transmembrane proteins.

For Gram-positive bacteria (Figure 1.4, left) the inner membrane is enveloped in a thick peptidoglycan layer. If its thickness in Gram-negative bacteria reaches a few nanometers, in Gram-positive ones it spans from 30 to 100 nm. This thick layer is threaded by long anionic polymers (the teichoic acids), mainly composed by glycerol phosphate, glucosyl phosphate, or ribitol phosphate repeats [Swoboda et al., 2009]. Disseminated in this layer there are several surface proteins with various functions, among which adhesins, which attach to components of the host extracellular matrix.

Gram-positive membranes are generally more permeable because they do not possess a double-membrane structure: the peptidoglycan layer is partially permeable to small molecules, however positively charged drugs are likely sequestered by the negative teichoic acids, challenging their penetration but promoting a disruptive action at the membrane level [Malanovic, 2016].

In addition to the aforementioned structures, many bacteria are coated with either the S-layer, made of glycoproteins, or polysaccharised capsules Silhavy et al. [2010]. These external envelopes, and especially the S-layer, mediates a broad range of functions, such as the cell wall formation, and, for pathogenic bacteria, influence the response of the host immune system to them [Fagan and Fairweather, 2014]. The presence of these layers increase the complexity of the bacterial membrane structure and thus the challenges a drug must overcome to be effective against pathogens.

Finally, it must be remembered that many proteins tessellate the bacterial membrane [Silhavy et al., 2010]. As such, the charge and fluidity of the membrane is not only determined by the lipids present but, especially in proximity of the interface with proteins, it can vary according to the identity and the local concentration of the pro-

teins.

Comparison with mammalian membrane The fact that AMPs tackle negatively charged membranes is crucial for their selectivity, i.e. that they are harmless for the mammalian cells they are produced from [Glukhov et al., 2005]. This is possible because mammalian cells have a different membrane composition. They present a single membrane, which is rich in proteins (up to 50% of its volume) and in lipids, and a small percentage of carbohydrates, mainly embedded in glycoproteins, which promote cell-cell recognition.

The lipidic component is abundant in zwitterionic phospholipids such as phosphatidylethanolamine, phosphatidylcholine, and sphingomyelin, providing a neutral net charge [Spector and Yorek, 1985; van Meer et al., 2008]. Furthermore, the mammalian cell membrane has a high content of cholesterol [Yeaman and Yount, 2003; Lai and Gallo, 2009], a sterol fat, which is proposed to stabilise the membrane ordering the lipids around itself and thus regulating the membrane fluidity across different physiological temperatures; it is also though to favour a better accommodation of the perturbations caused by AMPs [Zasloff, 2002]. Strictly speaking, some negatively charged lipids are present in a few mammal cell types, however they are located in the inner leaflet, while the zwitterionic phospholipids are more abundant in the outer leaflet, in an asymmetric composition [van Meer et al., 2008; Matsuzaki, 2009]. This structure promotes weaker interactions between AMPs and mammalian cells with respect to bacterial ones, as the former is driven mainly by hydrophobic interactions, while the latter by electrostatic ones. The interaction is useful to quantify the perturbation of AMPs on the membrane, and as such has been analysed in the present work.

Another relevant difference between bacterial and mammalian cells is that the first ones have typically a higher transmembrane potential - the difference of electrostatic potential between the inside and the outside environment. For bacteria it falls between -130 and -150 mV, while for mammalian cells between -90 and -110 mV [Yeaman and Yount, 2003; Matsuzaki, 2009; Ebenhan et al., 2014]. Given that a potential generates an electric field across the membrane, the higher this is, the higher the resulting electric field pointing from outside to inside the cell. A field in such

direction pushes cationic compounds on the outside of the membrane toward the membrane itself. Therefore a stronger bacterial transmembrane potential may promote an enhanced - and thus disruptive - interaction of AMPs with the cell, contributing to AMPs selectivity between bacteria versus mammals [Yeaman and Yount, 2003].

1.2.2 Common mechanisms of action of AMPs

Investigating the perturbation and disruption of a bacterial membrane by antimicrobial peptides is a key point of this work, therefore it is important to highlight the mechanisms known so far through which AMPs reach this outcome. As already mentioned, many AMPs have a positive charge which facilitates the binding to the membrane via charge-charge recognition; accordingly, Arginine and Lysine residues are usually abundant in AMPs sequences. However, the disruptive action takes place through the interaction of AMPs with the hydrophobic core of the membrane, therefore their sequences contain hydrophobic aromatic residues too, especially Tryptophan, which favours the anchoring to the lipid core [Chan et al., 2006]. Overall, AMPs resort often to adopt an amphiphatic structure to segregate the hydrophilic from the hydrophobic amino acids and thus to act at the interface between membrane and solution. It is interesting to notice that some of them fold into the active structure only nearby the membrane, where they expose their hydrophobic components to face its core, while in solution these ones are preferentially buried inside the peptide fold to be screened from the solvent [Nguyen et al., 2011]. Common folds adopted by AMPs are both α-helix or β-sheet rich structures. Amphiphatic α-helices present a charged side which is tailored to face towards the phospholipid head groups and an hydrophobic ones which is favourably buried into the acyl chains core. A similar arrangement is found for structures rich in β-sheets, such as β-hairpins (Figure 1.5).

Membrane disruption Several models have been proposed to describe the exact mechanisms of AMPs penetration after they bind to the cytoplasmatic membrane, and how their action leads to membrane permeabilization (Figure 1.3) [Brogden, 2005; Nguyen et al., 2011].

For a single copy of an amphiphatic helical AMP, the proposed mechanism of action

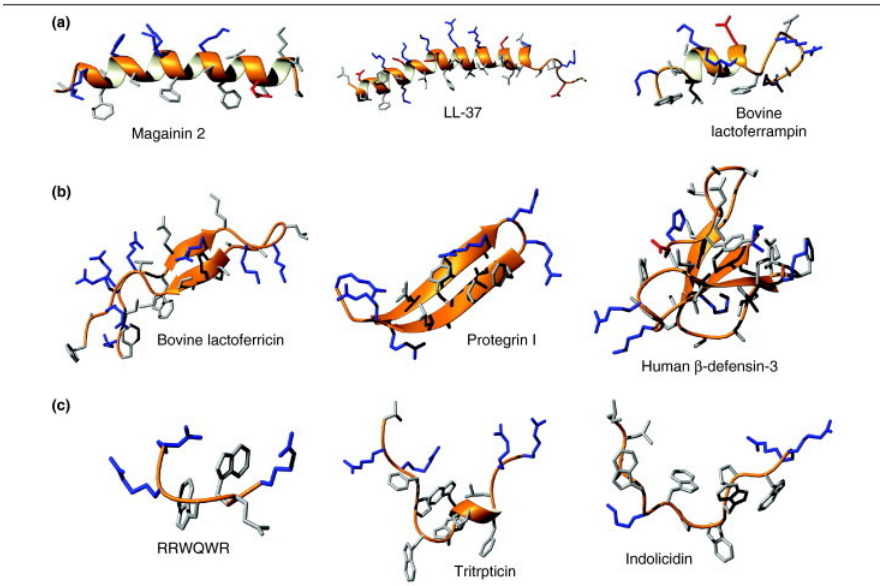


Figure 1.5: Structure of some known AMPs. (a) Helical structures, (b) β-sheet structures, (c) disordered ones. Reproduced from [Nguyen et al., 2011].

suggests that initially the peptide is attracted with its charged side to the membrane and lies parallel to its plane, with the hydrophobic side unfavourably exposed in solution. Then the helix rearranges to have the two faces in the respective favourable regions. Subsequently, the helix axis starts to form an angle with the membrane plane, and finally inserts deeper into the lipid core, often spanning the full membrane thickness [Ebenhan et al., 2014]. Similarly, for β-sheet rich structures, it is suggested that they insert within the membrane after a first flat approach. The final insertion arrangement depends on the peptide characteristics, length and specific structure, and the interactions with other copies of the peptide.

The picture becomes more complex for oligomer-mediated insertion, i.e. when the action is triggered by the combined action of many copies of the peptide. At low peptide to lipid ratio, the favourable configuration is represented by peptides lying parallel to the membrane plane as described previously [Yang et al., 2001]. An increase in peptide concentration triggers the transition to an inserted state: the organisation of AMPs inside the membrane core can assume different configurations, as described below [Brogden, 2005; Nguyen et al., 2011; Ebenhan et al., 2014; Mahlapuu et al., 2016] (see Figure 1.3).

The "barrel-stave" model proposes that AMPs insert perpendicularly into the bi-

layer. Recruitment of peptides in the same area results in the formation of a transmembrane pore with a central lumen. The walls of the pore are constituted by the hydrophilic face of the peptides, while their hydrophobic side is interacting with the lipid tails around the pore. This model is adopted, for example, by the α-helical AMP alamethicin, which forms voltage-dependent ion channels by aggregation of four to six molecules [Spaar et al., 2004].

In the "toroidal" pore model instead, the insertion of peptides forces the phospholipid to bend continuously from one leaflet to the other. The toroidal model differs from the barrel-stave model as the peptides are always associated with the lipid head groups even when they are perpendicularly inserted in the lipid bilayer. Toroidal pores are induced by α-helical magainins, protegrins and melittin [Yang et al., 2001; Matsuzaki et al., 1996; Hallock et al., 2003], and lead to membrane perturbation which extends further away from the pore than in the barrel-stave case. As a comparison, alamethicin induced barrel-stave pores have an inner and outer diameters of 1.8 nm and 4.0 nm respectively [Spaar et al., 2004], while magainin-induced toroidal pores have variable sizes, with an inner diameter of 3.0-5.0 nm and an outer diameter of 7.0-8.4 nm [Matsuzaki et al., 1997].

Finally, in the "carpet" model, the accumulation of AMPs on the surface of the membrane, laying parallel to it, causes tension in the bilayer: the membrane is disrupted by them in a detergent-like manner, leading to the formation of micelles. The critical threshold concentration triggers a cascade effect, as the first disruption allows the penetration of AMPs in the inner side of the bilayer. The cooperation between peptides on both sides of the lipid membrane enhances the AMP-induced curvature causing accelerated disruption. The "carpet" model mechanism is observed for peptides presenting an α-helical structure (like melittin [Ladokhin and White, 2001]) or for several helices connected by short loops (ovispirin [Yamaguchi et al., 2001]).

The prevalence of examples with an helical structure derives from the fact that the understanding of how helical AMPs function is often easier than the one of β-sheet rich structures. Indeed, helices have a well defined fold (at least nearby the membrane environment), a compact structure, and often a clear segregation of complementary patches that can attract other copies of the peptide and thus promote the self-assembly

process necessary for pore formation.

On the contrary, many β-sheet AMPs have a more flexible structure, and more diversified mechanisms of action [Nguyen et al., 2011; Mahlapuu et al., 2016]. AMPs rich in β-sheets can be divided into β-hairpins and peptides from the defensin family [Nguyen et al., 2011]. Many representative of the first class disrupt bacterial membranes via formation of toroidal pores: as an example, porcine peptide protegrin I assembles into a β-barrel structure when in contact with anionic membranes and triggers toroidal pore formation. Instead, it folds into β-sheet aggregates on the surface of cholesterol containing membranes, thus acting selectivity on bacterial membranes only [Tang and Hong, 2009].

In the case of defensins, many mechanisms are known according to the specific member of the family [Lehrer, 2004; Kagan et al., 1990; Takeuchi et al., 2004]. Although various descriptions of membrane damage have been reported, and include ion channels, transmembrane pores and extended rupture of the membrane, they are likely related, being a modulation of a similar acting principle.

The aforementioned models provide an important insight into how AMPs work. However, they are often obtained under artificial conditions, i.e. on model membranes. The composition of these membranes is necessarily simplified with respect to the true bacterial envelope, retaining often only one lipid layer made of few lipid species. Reducing the variety of lipids present in the model membrane is justified by the similar chemical characteristics lipids have (e.g. hydrophobic tails versus polar/charged head). However, a significant percent of the bacterial membrane is constituted by proteins, which can cover up to the 30% of the surface [Hsu et al., 2017]. The density of proteins on the surface and their large size is such that all the regions in the membrane are influenced by their presence, due to electrostatic interactions with the proteins and the consequent reduced lipids mobility, up to 2 nm away from it [Hsu et al., 2017]. The poration models presented before are obtained for bulk, unperturbed lipids: especially for the toroidal model, a large radius is required to allow the smooth transition necessary to create the toroidal shape. In the presence of proteins in biological concentrations, little to no region of bulk lipids is found. Thus, AMPs are must work in a different way and produce different pore shapes: whether the protein is stabilising or destabilising

the membrane depends on the specific case in exam, but it is not unlikely that the lipid-protein boundary constitutes a region of peptide insertion given the discontinuity present there.

In light of these considerations, poration models and poration experiments on model membranes are still important to compare the action of different AMPs and test novel ones, but might not explain the efficacy of AMPs in a complete fashion.

Alternative mechanisms of action Finally, many non-lytic mechanisms are suggested for AMPs, especially for β-sheet structures: defensin A from P. terramovae reduces the cytoplasmic potassium concentration [Brogden, 2005], partially depolarising the inner membrane. Tachyplesin from horseshoe crabs is able to bind to the minor groove of DNA, interfering DNAprotein interactions [Yonezawa et al., 1992]. Bovine lactoferricin can act synergistically with other antimicrobial agents by affecting the transmembrane potential and proton-motive force, resulting in inhibition of ATP-dependent multi-drug efflux pumps [Gifford et al., 2005]. Moreover, after translocation within the cell, bovine lactoferricin can also inhibit DNA, RNA and protein synthesis.

Section 1.5.1 will treat in detail the functioning of this AMP, distinguishing its role as membrane active peptide as opposed to intra-cellular targeting compound. Many works have focussed on locating the section of the sequence performing the membrane disruptive activity [Tomita et al., 1994; Schibli et al., 1999], to understand whether it retains the efficacy regardless of the fold. These type of investigations provides the discovery of minimal functioning antimicrobial blocks, which promotes the understanding of how AMPs work in general, and boost the design of synthetic AMPs tailored for specific functions.

1.2.3 Mechanisms of resistance to AMPs

Antimicrobial peptides were introduced here as a class of new drugs and a possible solution to the crisis of antimicrobial resistance. Any new drug entering the pool of the clinically approved compounds is (at least temporary) a solution to the problem of resistance; but it must be clarified that bacteria can develop resistance to AMPs too. Nevertheless, this is generally not based on dedicated genes that are conferred

by horizontal gene transfer, as in the case of many antibiotics resistance mechanisms [Peschel and Sahl, 2006]. Because of that, a certain increase in bacterial resilience after exposure to the peptidic drug is to be expected, but it is less likely to spread quickly to other species.

Some of the mechanisms of AMPs resistance are similar to the ones employed by bacteria to counteract small molecule drugs, for example over-expression of efflux pumps to dispose of AMPs, degradation of the peptide by extracellular enzymes and sequestration by the bacterial or biofilm matrix to prevent accession to the target [Peschel and Sahl, 2006].

Differently with respect to antibiotics hydrolysis, AMPs proteolitic degradation is operated by proteases, secreted on the extracellular side of the membrane specifically to destroy other proteins. Linear AMP are more prone to this type of degradation [Sieprawska-Lupa et al., 2004], as opposed to the ones presenting disulfide bonds [Peschel and Sahl, 2006], such as defensins, which nevertheless can be hydrolysed by more specific enzymes [Nelson et al., 2011].

But the most specific mechanism of AMPs resistance concerns modifications of the bacterial cell envelope: bacteria modify the characteristics of their surface to prevent the efficient binding of an AMP, even in the eventuality that the peptide reaches the bacterial envelope intact. The target of such modifications are different for Gram-positive and Gram-negative bacteria, according to their distinct cell envelopes. for example, Gram-positive bacteria change the structure of their teichoic acids (TA): D-Alanylation of TA observed in *S. Aureus* adds a positive charge to it, reducing the attraction of cationic AMPs and in turn increasing the cell wall density, so reducing the surface permeability [Saar-Dover et al., 2012].

In Gram-negative bacteria a positive charge can be added to lipopolysaccharides (LPS) by addition of amine-containing molecules [Moskowitz et al., 2004] or by removing phosphate lipids (which have a negative charge) from lipid A [Wang et al., 2006b]. Moreover, the cytoplasmic membrane can be modified as well, as this is the final target of many antimicrobial peptides: in the eventuality that AMPs successfully reach this membrane, they are attracted to its surface by the negative charge of the lipids composing it, in particular phosphatidylglycerol (PG) and diphosphatidylglycerol (DPG, also

called cardiolipin). Their negative charge can be masked by amino-acylation of the PG head group, so that the final compound repels AMPs through electrostatic interaction [Peschel et al., 2001]; alternatively the overall rigidity of the cytoplasmic membrane can be enhanced, by an increase in saturated acyl chains which has been proven to confer resistance [Kumariya et al., 2015]. To be noticed that resistant bacteria often employ many of the aforementioned strategies at the same time [Band and Weiss, 2014].

1.2.4 Principles of AMP design

The study and classification of AMPs provide knowledge on the characteristics a sequence must have to perform an antimicrobial function. As discussed in Section 1.2.2, there are some features which, comprehensively, help in discriminating AMPs against non antimicrobial peptides. The constantly increasing amount of data available is gathered in several curated databases (APD3 [Wang et al., 2016a], DBAASP2 [Pirtskhalava et al., 2016], dbAMP [Jhong et al., 2019], antiBP2 [Lata et al., 2010], amPEP [Bhadra et al., 2018]), which catalogue AMPs (or subclasses of them, like membrane active, biofilm active or haemolytic peptides) based on such features.

To recapitulate, AMPs are charged moieties, usually cationic, and their potency, but also their haemolytic activity, is often related to their net charge [Jiang et al., 2011]. They can assume both α-helical and β-sheet fold; which helps creating an amphiphatic structure, to accommodate both the charged residues mentioned and the hydrophobic ones they possess as well. Finally, they need good solubility to prevent aggregation in the aqueous environment before reaching the target.

The knowledge of AMPs sequence-activity and structure-activity relationships is beneficial to design new peptides with improved characteristics. In particular, improved specificity against bacterial species; stability against the action of proteases (allowing a longer residence time in the body); and low cytotoxicity at the therapeutic dose required. The need for such improved peptides lies in the fact that natural AMPs constitute a first broad spectrum defence that our body employs against infectious bacteria, and thus they are often of mild potency. However, foreseeing their application as future drugs, it is desirable to tailor them to fulfil different criteria according to the infection to treat. At the present state of the art, a golden rule for the design of such

sequences is still missing, however several methodological approaches to AMP design have been explored, and they can be grouped in three main lines: template based studies, biophysical studies and virtual screenings [Fjell et al., 2011].

Template based studies The main idea behind template based methods consists in modifying existing antimicrobial sequences in the direction of the desired characteristics. As such, the effort is restricted on a subset of all possible sequences. The most widely explored templates are helical peptides, because several of them (cecropin, magainin and protegrin) have been well characterised [Wang et al., 2015a].

Alanine scanning [Migoń et al., 2018] and all amino acids scanning [Hilpert et al., 2005] performed for every residue in a sequence, provide information on the role of each of them, pointing at the most suitable mutations. High-throughput methods for synthesis and characterisation allow nowadays for such thorough investigation in the case of short AMPs [Hilpert et al., 2005]. However, these methods focus on single amino acids and can not take into account the interplay between residues, nor the three dimensional structure of the peptide. Because of this, often the results of these studies cannot be generalised to other sequences, hence the recent effort to integrate structural information on template based models [Liu et al., 2018; Jiang et al., 2011].

A complementary approach to single point mutations on known peptides consists in designing minimal antimicrobial blocks: synthetic AMPs have been produced with only Lysines-Leucine, or Arginine-Valine combinations to produce amphipathic helices [Deslouches et al., 2005]. Text based models where amino acids constitute the letters and patterns occurring in natural AMPs are the grammar rules are trying to capitalise this findings into general rules [Loose et al., 2006; Cipcigan et al., 2018; Spänig and Heider, 2019].

Biophysical studies Biophysical studies aim at understanding the functioning of AMPs investigating their structure. Free energy perturbation, Molecular Dynamics (MD) simulations and thermodynamics calculations can all provide knowledge on how the three dimensional arrangement of residues is important to allow their functional role. These techniques give an insight into the mechanism of action of an AMP but

their drawback lays in the high computational cost, preventing the reproduction of phenomena of the order of millisecond and a systematic study of many sequences at once. A detailed overview of the state of the art, advantages and drawback of MD simulations will be given in Chapter 2.

The strength of biophysical studies lay primarily in the fact that they exploit the whole information available on a system (sequence, structure, chemistry). Thus, they can single out the interactions that are crucial for a mechanism, clarifying whether they can be transferred to a different environment. In this respect, they provide a generalisable knowledge applicable to different systems and to the design of novel AMPs at the atomistic level. Extensive examples of this workflow are given in Section 2.4.

Virtual screenings Contrary to biophysical assays, virtual screening methods are employed to analyse a large number of sequences, when an experimental or computational test of all of them would be prohibitive. The concept of these methods consists in identifying descriptors which allow to predict the potency of the sequence: from the analysis of a database of AMP with known activity, a model is created and used to score novel synthetic sequences [Fjell et al., 2011; Kleandrova et al., 2016].

The recent evolution of machine learning (ML) techniques, and artificial neural network in particular, gave a great impulse to virtual screening of AMPs (for a historically informed review see [Fjell et al., 2011; Veltri et al., 2018]). Machine learning appears particularly suitable to the task as the potency of AMPs is determined by the combination of many factors, the relative weight of which can be difficult to identify. Moreover, it can help in the identification of relevant features traditionally overlooked.

Machine learning algorithms require a consistent pool of data to be trained in a satisfactory way. However, modern high-throughput synthesis methods, together with surrogate measures of bacterial killing, are allowing to assessed the antimicrobial properties of thousands of sequences at once [Cherkasov et al., 2009]. This is a first step toward an automated and general procedure for AMP design.

1.2.5 Clinical applications

Antimicrobial peptides have been studied for many years, however the push to capitalise them has been delayed by many factors, including production costs, and lack of interest in the face of more potent small molecules. The constant increase of AM resistance has resulted in more effort focusing on AMPs, mainly from small biopharmaceutical companies, and at present several of these preparations are in clinical trials, either in phase 1 or 2 [Naafs, 2018].

The two major problems encountered so far for AMPs sequences in trial are the liability to proteolytic degradation, and the unknown toxicology profile when administered systemically [Hancock and Sahl, 2006]. For the last reason in particular, many of them are in trial for topical use only, as they are deemed unsuitable for internal administration. Design of novel AMPs can be tailored to improve the liability to degradation, for example introducing D-amino acids, non natural amino acid analogues of opposite chirality, which, with appropriate formulations, are mimetic to the immune system [Wipf et al., 2009]. Moreover, machine learning protocols can help in pre-screening their toxicity through virtual screening methods.

Overall, antimicrobial peptides remain a promising tool to counteract infections and, as their design is still - comparatively - in its infancy, there is room to explore novel applications and synthesise improved sequences apt to get to the clinical stage.

1.3 Gene therapy

Alongside the new compounds used to counteract bacterial infections, we want to bring the reader's attention to another class of therapies which is relevant for the work of this thesis and has been developed in the last decades for the treatment of non infectious diseases: gene therapy. In recent years it has greatly evolved and gained attention for the treatment of tumours, genetic diseases and complex acquired disorder [Anguela and High, 2019]. The key concept is the delivery of genetic material to cells of diseased state which possess a faulty copy of a gene, to influence its expression. Such fault can result in lack of synthesis of the protein of interest or in its misfold and/or misfunction. The correction can be performed in three different ways (Figure 1.6) [Anguela and High,

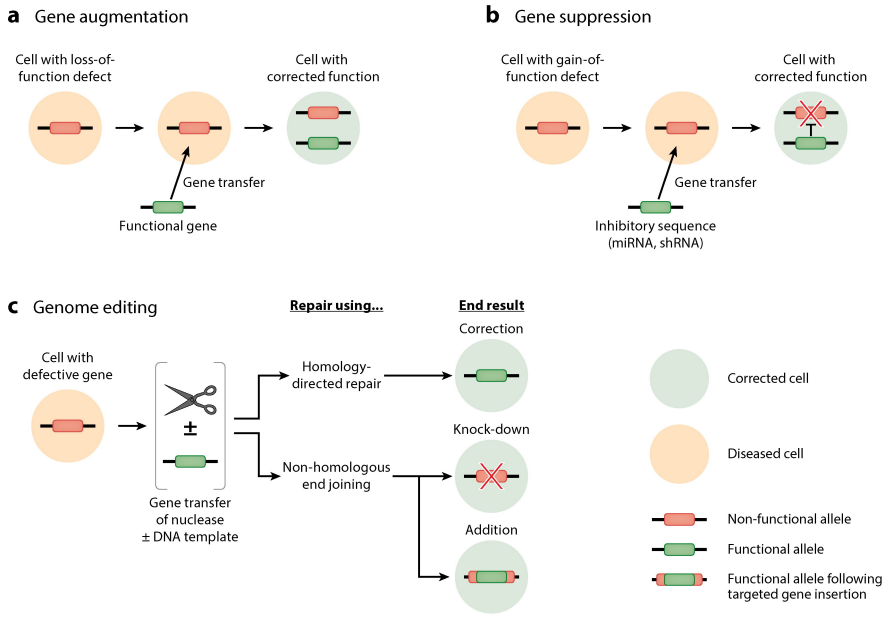


Figure 1.6: Principles of gene therapy. Reproduced from Anguela and High [2019].

2019]: introducing an healthy gene copy to restore the normal functionalities of the protein of interest; suppressing a detrimental gene (particularly useful in the case of cancer, to impede cancer cells replication); or correcting base pairs mutations to restore the original healthy sequence (gene editing).

Despite the challenges posed by the development of genome editing tools, and the risk associated to them (for example the possibility of deleterious insertional mutagenesis or deleterious immune responses), at present six gene therapies have received approval in the Western world [Anguela and High, 2019], with many more undergoing regulatory review.

One of the main problems in the development of such therapies lies in the identification of a suitable vector for the genetic material, either DNA, microRNA (miRNA) or small interfering RNA (siRNA). Delivery of free genome in solution results in poor internalisation and low therapeutic effect. Thus nowadays the outlook of gene therapy research lies not only in improving specific cargos, but also in the research of appropriate vectors with low toxicity, low induced immune response and high delivery efficiency. Viruses can be used, modifying their genome to include the necessary sequence and remove the ones promoting viral replication [Naldini, 2011; Mingozzi and High, 2011] (see Section 1.4.1), but synthetic vectors are now investigated for a virus-free delivery

strategy. The system studied in this thesis proposes, among its other functions, to deliver genetic material into human cells.

1.4 Delivery of therapeutic material

The problem of gene delivery sets a parallel with the small drug one, introduced at the beginning of the chapter. Indeed small molecules need delivering agents to be efficiently internalised in the cells, in the same manner that genes do, and it is known that some barrier (such as the Blood Brain one) are more selective in the classes of molecules they allow to permeate [Pattni and Torchilin, 2015; Krol, 2012].

To reach the aimed organ, therapeutic molecules must be compatible with the different cellular environments they cross, but be preferentially retained, and act only on the ones they are designed for. This implies a subtle balance between an invasive activity on one side, and mimesis on the other, to minimise the possibility that the compound is recognised as dangerous and disposed by the immune and reticuloendothelial systems.

For the above reasons, research has focused on developing systems to assist the delivery of drugs. A mimetic carrier can not only improve delivery, but also be designed to selectively bind to particular tissues, or to trigger the drug release after a delay in time, or only upon changes in environmental variables (for example pH), to reduce drug concentration in non targeted regions Pattni and Torchilin [2015]. A stand alone field of research has then focused on the development of delivery vehicles irrespective from the quest for new drugs. The optimised products of the two separate efforts can then be paired according to the condition to address, to give a successful therapy.

At present, both organic and inorganic molecules have been employed to build drug vehicles, to offer a range of different physico-chemical characteristics useful to target different cells [Hughes, 2005]. Inorganic materials, such as metal nanoparticles or carbon nanotubes, are in experimental phases and hold promise for promoting sustained drug release [Boisselier and Astruc, 2009; Depan et al., 2011], while inorganic polymers have been employed extensively and since a long time as drug eccipients [Lammers et al., 2009; Liechty et al., 2010; Nicolas et al., 2013].

The use of organic molecules instead aims at mimicking the materials present in the body, such as lipids, DNA and peptides, in an effort to reduce toxicity and favour the internalisation of the vehicle in the target cells. In particular, lipids have been widely employed for the delivery of both soluble and insoluble drugs thanks to their amphiphatic structure [Pattni et al., 2015; Jain and Pillai, 2017; Yingchoncharoen et al., 2016; Bunker et al., 2016]. DNA scaffolds instead are a novel tool still at the experimental stage which have been proven successful in delivering anticancer agents [Zhang et al., 2014; Jiang et al., 2012].

But peptides are the focus of this work, and these molecules prove once more they can have functions beyond the ones they are naturally designed for. In the next paragraph we elucidate some rationale and examples of how peptide can be tuned for assembly and thus for the needs of drug delivery.

1.4.1 Peptidic scaffolds

A widely used and trustworthy mimetic vehicle comes, quite surprisingly, from the world of pathogens: viruses have co-evolved with humans, to be able to penetrate into cells where they complete their reproductive cycle [Lobo et al., 2009]. Therefore their capsid, the peptidic shell encapsulating the viral genome, is highly suitable for cell penetration. The first application sought historically was to employ genome-free viruses to stimulate the natural immune response against the respective genome-loaded ones, creating viral vaccines - similarly to how inoculation of dead bacteria counteracts their infections [Lauer et al., 2017]. Later in the history, their potential as cargo carrier was pursued modifying the original genetic material to include sequences beneficial for the host cell, and inactivate their duplication [Daya and Berns, 2008]. Since then, many efforts have focused on synthesising in vitro gene-free capsids, either as they appear in nature [Wu et al., 2009] or designing artificial building blocks which assemble in so called Virus-Like particles (VLPs) [Schoonen and van Hest, 2014] aimed at triggering a lower immune response (Figure 1.1, i). Similarly to other delivery vehicles, the surface of VLPs can be functionalised with additional molecules to improve the target selectivity and biocompatibility, while the peptidic scaffold grants robustness to the structure. Finally, VLPs loaded with drugs can be tuned for an efficient intra cellular

release [Ma et al., 2012].

A step further in engineering peptidic structures is the design of self-assembling functional blocks from first principles. Indeed, self-assembling peptides can form nanostructures ranging from nanoparticles to nanotubes, nanofibers, nanorods and hydrogels [Fan et al., 2017]. The variety of amino acid available makes peptidic structures compatible with both hydrophilic and hydrophobic drugs, according to their amino acid composition [Ma et al., 2012]. The peptidic self-assembly is modulated by the peptide length and its hydrophobic or hydrophilic character, given by its amino acid composition: simple phenylalanine dipeptides (Figure 1.1, j), were shown to assemble in a multi-scale process into nanotubes able to load drug molecules [Silva et al., 2013]. The relatively small diphenylalanine building block is nevertheless complex as it bears two charged termini (as the process is observed at neutral pH), and two aromatic hydrophobic rings, so that the dipeptide is driven towards assembly by the hydrophobic forces acting on the phenylalanine side chains and the complementary charges of the termini.

In a different approach, longer sequences (which organise spatially in well studied α-helical or β-sheet secondary structures) can be employed to guide the assembly at the tertiary structure level (see some examples in Figure 1.5). This knowledge is possible as proteins are a fundamental component of the human body and as such an updated database of their structure is available (the Protein Data Bank [Berman et al., 2000]) and can be queried to understand how small peptides hierarchically assemble into larger units. Finally, the vast literature on their interactions with membranes, cell receptors and in general biological components, can inspire the design of building blocks sensible to particular triggers within the body. From this background, the outlook of protein design often goes in the direction of synthesising exotic, non natural geometries for multifunctional materials [Yeates, 2019; Malay et al., 2019].

1.5 Closing the circle: an antimicrobial drug delivery vehicle

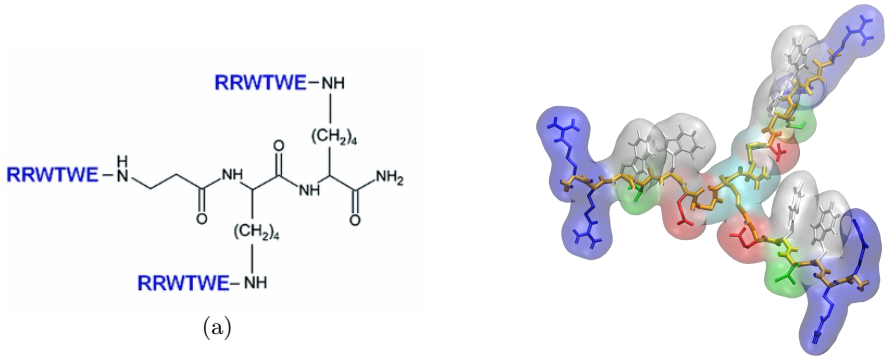
Twice in this introduction peptide design has been brought to the reader's attention. First, it can produce antimicrobial peptides with improved potency and/or selectivity, and/or reduced toxicity. Second, it can engineer self-assembling building blocks for the

formation of delivery scaffolds. As design is not bound to natural rules, it can foresee and imagine multifunctional materials which are not observed in nature. In particular, the question arises whether it is possible to engineer peptides able to perform both an antimicrobial and a delivery function at once.

Self-assembling antimicrobial compounds would have a twofold interest for medical applications. First of all, the assembly is functional to the antimicrobial activity: many AMP sequences have a weak potency, and only a high (critical) concentration can trigger the bactericidal mechanism (e.g. the insertion into the membrane or its lysis, see Section 1.2.2). Second, the assembly can perform additional delivery functions, if it is able to either organise in a tailored structure (for example a capsule able to host a drug), or to co-assemble with the cargo of interest.

Out of all the possible applications, an ideal scope is perhaps the use of such vehicles to deliver drugs to treat metabolic or genetic diseases: while the cargo tackles a defect of the host system, the vehicle can counteract the proliferation of bacteria. This is particularly important when the host immune response is weakened and normally harmless infections can spread and cause damage. As such, the cargo is not bound to be a small molecule, as long as it can effectively co-assemble with the peptidic carrier. As mentioned in the previous section, gene therapy is also an actively expanding field which looks with interest at the development of vehicles. Given that viruses have been the first choice for DNA/RNA delivery so far, peptidic carriers seem their natural evolution.

Given the above premises, it is evident the importance of pursuing the research on novel multifunctional peptidic materials. As mentioned when discussing AMPs design, to better understand these systems, each of them must be characterised by itself, as a generalised knowledge is still lacking. With this aim, this thesis proposes to elucidate the behaviour of a specific synthetic self-assembling peptide, suitable for antimicrobial activity and gene delivery strategies. Its characterisation will complete the knowledge on its mechanisms of action and complement the general information already known on the class of such functional building blocks. This is crucial to engineer new synthetic blocks with improved characteristics, either regard their antimicrobial activity, assembly performances, or tailored cargo delivery.



(a) (b)

Figure 1.7: Capzip molecule: (a) scheme of chemical formulation, reproduced from Castelletto et al. [2016]; (b) 3D bonds representation: backbone in yellow, amino acids side chains colored by residue type (blue positive, green hydrophilic, white hydrophobic) [VMD software Humphrey et al. [1996]].

1.5.1 The capzip molecule

The molecule capzip [Castelletto et al., 2016] (Figure 1.7) has been designed to perform the functions mentioned above at once. To recapitulate, the properties it must possess are:

1. assembly into nanoscale virus-like capsules with and without nucleic acids. This ensures that the vector can autonomously form and thus there is flexibility in the choice of the cargo;

2. antimicrobial activity of the molecule itself and of the capsule on a time scale useful for therapeutic applications;

3. promotion of gene transfer into mammalian cells when the peptide is co-assembled with the RNA strands, without causing cytotoxic and haemolytic effects.

Its design aims at building a template structure of minimal complexity, in order to synthesise only a short sequence. Arguably, short sequences are flexible in their assembly, as they do not rearrange into defined secondary structure: it is thus even more important to understand them and prove whether also small blocks can form ordered architectures. To satisfy the short sequence criterion and the required properties of the assembly, two design principles emerged: first the employment of a non-linear structure,

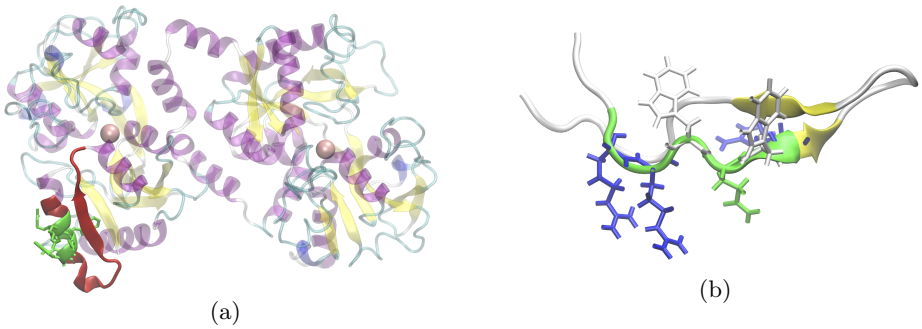
and second the use of a template antimicrobial sequence which is short and has proved antimicrobial efficacy.

There is indeed some evidence suggesting that the branching of short peptidic sequences without secondary structure tunes their three dimensional assembly [Gudlur et al., 2012; Breger et al., 2017; Zhao et al., 2018b]. Indeed, linear short structures can pack parallel one next to the other (as amyloids), while branching prevent this tight structure and favours more open ones. Moreover, many viral structures have a 3- or 5-fold symmetry at each of their vertices [Schoonen and van Hest, 2014]: reproducing this, albeit in a different structure, is likely to improve its assembly properties. For these reasons, a short peptidic scaffold constituted by a β-Alanine and two Lysins which can host other peptidic branches has been engineered as core of the capzip molecule.

Regarding the second principle, the antimicrobial sequence selected has been derived from the AMP bovine lactoferricin, which is in turn a portion of the lactoferrin protein, and is six amino acids long (see next paragraph). Three copies of it are covalently bonded to the N-terminus of the scaffold sequence and to the nitrogen atom of the Lysin side chains (Figure 1.7(a)). As AMPs are usually cationic, and indeed the overall structure has a $+6e$ charge at physiological pH, the co-assembly with anionic RNA sequences is arguably naturally inherited by the molecule.

Lactoferrin Lactoferrin is a 80 kDA iron binding protein present in milk (in which it is most abundant, hence its name), saliva and other secretions, as well as in leukocytes (Figure 1.8(a) shows the bovine homologue of the human protein). It works as an iron binder and provides a natural defence against bacteria and fungi [Sánchez et al., 1992; Arnold et al., 1977, 1980; Kirkpatrick et al., 1971; Jahani et al., 2015], constituting a first defence for infants.

Lactoferrin contributes to bacterial suppression in several ways. At present, its known modes of action fall in three categories: first, thanks to its iron sequestering capabilities, it removes essential substrate required for bacterial growth [Farnaud and Evans, 2003]; second, it is implicated in the stimulation of different immunological cells (killer cells [Shau et al., 1992], polymorphonuclear leukocytes, and macrophages [Gahr et al., 1991]); finally it interacts with bacterial membranes and binds to the lipopolysac-



(a) (b)

Figure 1.8: (a) Bovine lactoferrin protein (PDB code 1BLF): transparent cartoon representation, colored by secondary structure. In red is highlighted the portion of sequence corresponding to lactoferricin. In green, the 6 amino acid antimicrobial sequence. In pink spheres the iron ions. (b) Bovine lactoferricin protein (PDB code 1LFC); cartoon representation, colored by secondary structure. In green, and in bond representation color coded by amino acid type, the 6 amino acid antimicrobial sequence. [VMD software Humphrey et al. [1996]]

charides of the bacterial wall, oxidising them and affecting the membrane permeability with consequent cell lysis [Farnaud and Evans, 2003]. The peptide fragment responsible for binding lactoferrin to the bacterial membrane, named lactoferricin (Lfcin), has been identified near its N-terminus and found to have a more potent bactericidal effect than intact lactoferrin on a wide range of bacteria [Gifford et al., 2005; Bellamy et al., 1992; Tomita et al., 1994; Wakabayashi et al., 1996] (Figure 1.8(b) shows the bovine homologue of human Lfcin). Similarly, an even shorter subsequence of human lactoferricin has been proven effective against bacteria as it depolarises the cytoplasmic membrane decreasing the pH gradient [Aguilera et al., 1999].

The bovine homologue of lactoferricin (LfcinB, Figure 1.8(b), PDB code 1LFC) has often a higher bactericidal potency than its human counterpart [Cochran et al., 2001] and therefore has been more extensively studied. It is a 25-amino acid sequence which adopts a helical conformation in the full structure but, once isolated, crystallises in a β-hairpin with a disulfite bridge nearby the terminals which stabilises the fold, but was shown to be not essential for the bactericidal activity [Cochran et al., 2001]. In solution, it adopts a flexible conformation, as assessed by NMR experiments [Hwang et al., 1998]. Further experiments on LfcinB subsequences identified a shorter antimicrobial core, constituted by the six amino acids RRWQWR [Schibli et al., 1999]. This core presents a characteristic Tryptophan zipper motif WQW, which appears very often in nature in

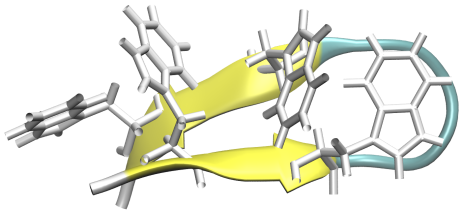


Figure 1.9: Example of Tryptophan zipper hairpin (PDB code 1LE0). Cartoon representation, colored by secondary structure, and bond representation for Tryptophan residues. [VMD software Humphrey et al. [1996]]

β-turns and β-sheets, paired to another copy of the same motif [Cochran et al., 2001] so that Tryptophan rings from facing strands are packed tightly against each other in an alternated way (Figure 1.9).

The six amino acid sequence contains both charged and hydrophobic residues, in line with the usual composition of antimicrobial peptides. Accordingly, its antimicrobial action is likely derived from the interaction with biological membranes through charge recognition first and aromatic rings insertion in a second moment.

To further elucidate this mechanism, several experimental investigations have been carried on, both on LfcinB and its subsequences. The binding of its antimicrobial core to sodium dodecyl-sulfate micelles was studied [Schibli et al., 1999], suggesting a favourable interaction of the aromatic residues with the micelles surface. Similar experiments were performed on large unilamellar vesicles, constituted by lipids modelling biological membranes [Nguyen et al., 2005]: ePE:ePC was chosen as a model of a mammal membrane, and ePE:ePG or ePC:ePG for a bacterial one (ePE egg α-phosphatidylamine, ePC egg α-phosphatidylcholine, ePG egg α-phosphatidylglycerol, with ePG anionic). The experiments showed preferential binding to the latter ones, based on Tryptophan fluorescence, suggesting a selective antimicrobial action on anionic membranes. Additional experiments performed on other LfcinB subsequences, some of which introduced mutations [Tsutsumi et al., 2012; Arseneault et al., 2010], investigated the binding to different model membranes. However, as both the systems and the experimental conditions are slightly different each time (in terms of length of

the original sequence and the type of membranes used), it is difficult to give a unified interpretation of the modes of action of lactoferricin-derived peptides.

Finally, an alanine scanning has attempted to clarify the role of each amino acid in the antimicrobial activity of the original 25 amino acid LfcinB peptide [Strøm et al., 2002]. The results suggested a membrane binding function for the Tryptophan residues, in line with one of the roles Tryptophan assumes in antimicrobial sequences [Chan et al., 2006]. Another possible role, however, involves its propensity to form hydrogen bonds, in which case the residue would position itself at the interface between solution and membrane, rather than inside the latter, which happens when the Tryptophan residue has the function of binding the peptide to the membrane [Chan et al., 2006].

The designed block From the active core of LfcinB (of sequence RRWQWR), a mutated sequence was obtained to comply the design criteria of a self-assembling building block. Two mutations were introduced to favour the assembly of arms belonging to different molecules in an antiparallel fashion. Specifically, given that the original sequence is found in a β-sheet (at least in the crystal lattice) facing a non-homologous one, the mutations aim at promoting β-sheet formation when two pairs of the *same* sequence are in proximity. Therefore, the Glutamine residue and the C-terminal Arginine of the lactoferrin motif were replaced with Threonine and Glutamic acid residues to have a self-complementary sequence RRWTWE: the pairing is promoted by the attraction of opposite charges at the ends of the sequence. Three copies of this sequence were thus covalently bonded to the scaffold as described previously (Figure 1.7(a)), to obtain a self-assembling molecule hosting multiple copies of an antibacterial sequence.

An additional layer of complexity has been added in the experimental investigation by synthesising a capzip molecule formed by D-amino acids only. The rationale behind this choice lays in the enhanced stability of D-peptides against proteolysis and their possibly low immunogenicity [Uppalapati et al., 2016; Arranz-Gibert et al., 2018; King et al., 1994], while the fold of the final molecule remains similar to its L- version. Moreover, as antimicrobial peptides bind to membrane rather than dock to a specific protein, their D-epimer is likely to retain the activity of the L- counterpart [King et al., 1994; Bland et al., 2001]. This approach opens a new landscape in the field of peptide

design, as it virtually multiplies the space of sequences available by a factor 2^N, with N the length of the sequence selected.

1.5.2 A viable systems: experimental background and question

A set of experiments has been performed to verify that capzip had the characteristics it was designed for. The results obtained on the molecule have been published in Castelletto et al. [2016], while more recent investigations extended and consolidated the previous findings [Kepiro et al., 2019].

Experimental results First, the assembly ability has been tested: the peptide (L-version) did not assemble in pure water (as verified by Dynamic Light Scattering), likely because of charge-charge repulsion, while in biological buffer (MOPS, 10 mM) at physiological pH of 7.4, it formed capsules with dominating size range of 20-200 nm. This was confirmed by images of the capsules obtained with multiple techniques, namely transmission electron microscopy (TEM), atomic force microscopy (AFM), and cryo-scanning electron microscopy (SEM) (Figure 1.10(a)). Analogous and more systematic investigation on D-capzip brought to similar conclusions (Figure 1.10(b)), restricting the dominant size range from 20 to 40 nm of diameter. Interestingly, some of the images collected showed ring structures, likely derived from collapsed capsules, suggesting they are hollow. Fluorescence microscopy of a capsule assembled from fluorescein labelled capzip confirmed this hypothesis, showing that the signal was coming from the walls of the structure only.

The fine structure of these assemblies appeared irregular to the resolution of the miscoscopy techniques employed and could not be fully determined. Some insight was given by Circular Dichroism (CD) spectra, which showed a profile characteristic of β-turns and contained elements of a β-sheet structure and of indole rings (minima at $\lambda \sim$ 200 nm and 214 nm) (Figure 1.11).

The assembly process was also tested and monitored in combination with small interfering RNA (siRNA) sequences. The co-assembly of a 21 base pairs double strand siRNA sequence with the peptide showed the formation of structures similar to the ones with peptide only: CD spectra highlighted the helical signal from RNA together

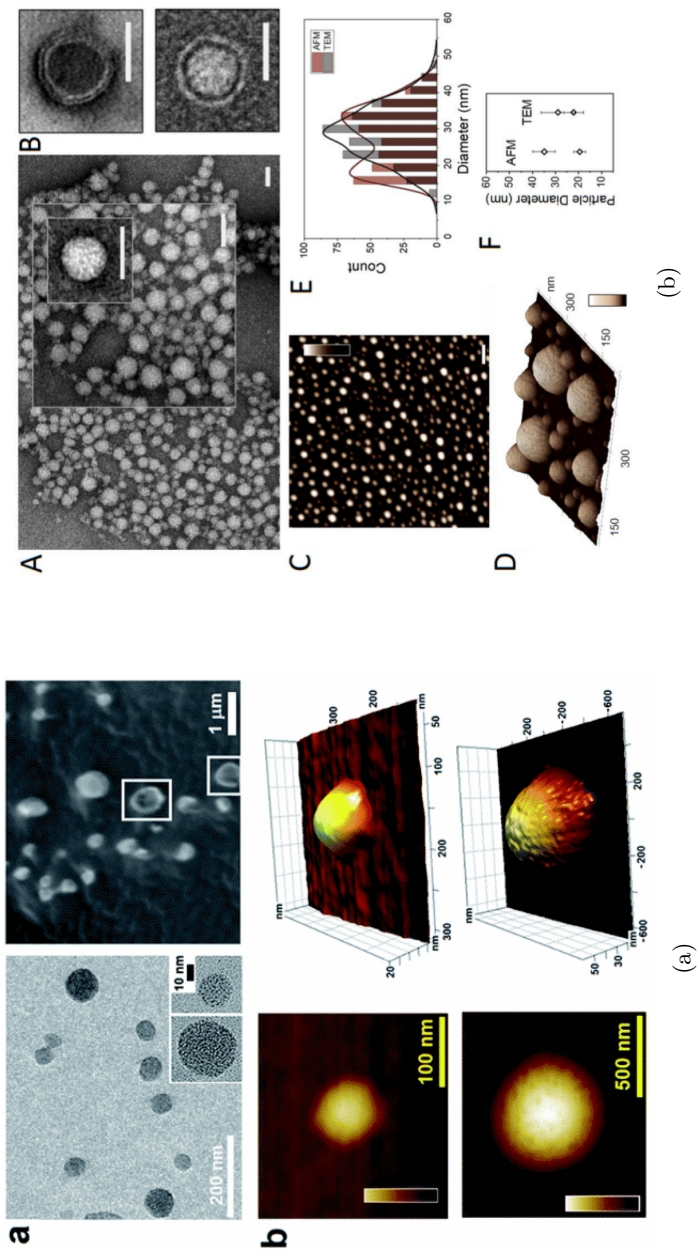


Figure 1.10: (a) TEM (top, left) and cryo-SEM (top, right) images of assembled capzip capsules. Collapsed ones are highlighted by white squares. In-air AFM topography of capsules and their 3D representations (middle - colour bar 20 nm and bottom - colour bar 60 nm). Reproduced from Castelletto et al. [2016]. (b) A: TEM images of assembled D-capsids (scale bars 50 nm). B: higher resolution TEM images of individual collapsed capsids (scale bars 50 nm). C: topography images of D-capsids obtained on a mica substrate by in liquid AFM (colour/height - and scale bars are 60 nm and 200 nm, respectively). D: 3D representation of D-capsids (height bar 65 nm). E: size distributions and dominating sizes of D-capsids by AFM and TEM. Error bars denote standard deviations. F: average sizes of dominating populations of D-capsids. Reproduced from Kepiro et al. [2019]. For both (a) and (b), assembly conditions: 100 M peptide, pH 7.4, 10 mM MOPS, 20 °C, overnight (15 hours).

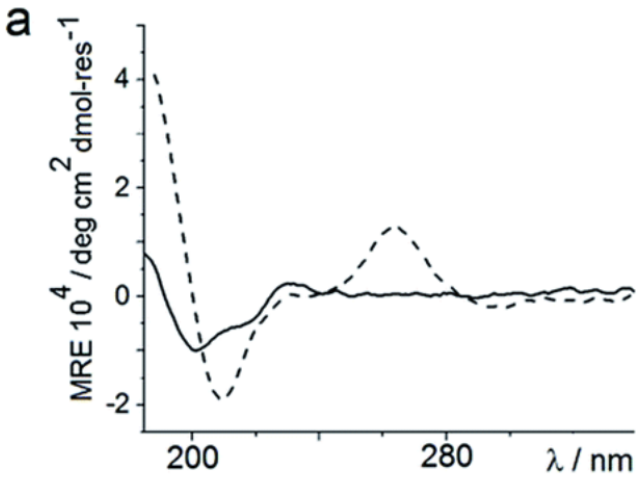


Figure 1.11: Circular Dischroism spectra for capzip capsules (solid line) and capzip with siRNA (30 μM, dashed line). Reproduced from Castelletto et al. [2016]

with the features proper of the peptide (Figure 1.11, dashed line). These co-assembled structures were tested for siRNA delivery in HeLa cells, using a fluorescent Alexa-labelled siRNA (Eurogentec, UK). The experiment showed that the internalisation occurred within the first hours from the transfection, suggesting an endocytic uptake (Figure 1.12, a). The presence of the peptide enhanced the internalisation with respect to the one obtained from a pure siRNA control (Figure 1.12, a), and was comparable with the uptake obtained with the standard transfection agent Lipofectamine®. Cell viability was assessed monitoring the level of green fluorescence protein expressed by the HeLa cell line chosen, which was stable over the first five hours of incubation with capzip, after which the fluorescence signal decayed.

Similar transfection experiments the fluorescent siRNA combined with D-capzip were assessed through flow cytometry assays. This method counts the number of fluorescent cells and can process thousands of them at once, giving large scale information on the efficacy of the transfection. Such assays confirmed that also D-capzip enhance the siRNA uptake in measure comparable with Lipofectamine® performances (unpublished results).

To further quantify the level of RNA internalisation, a micro RNA (mRNA) knock-down experiment was performed on a HeLa cell line with two housekeeping genes, ACTB (β-actin, targeted) and GAPDH (reference). The silencing of β-actin mRNA

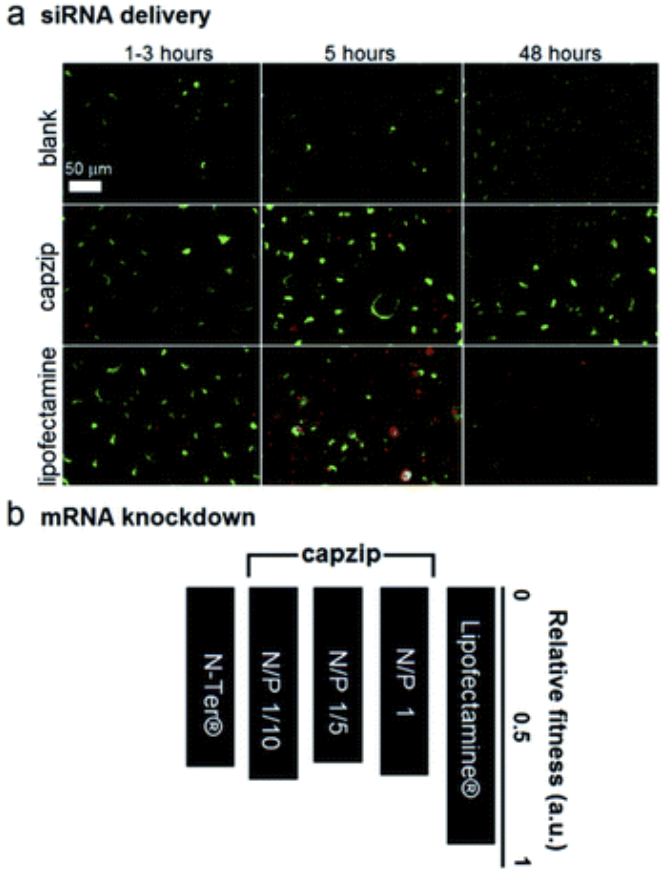


Figure 1.12: a: fluorescence micrographs of HeLa cells expressing green fluorescent protein used as internal background fluorescence for Alexa647-labelled siRNA, at 1/5 N/P (neutral to positive) molar ratios (green: GFP, red: siRNA). b: knockdown fitness of capzip and commercial Lipofectamine® or N-TER® (positive controls) normalised against siRNA alone (negative control) and the total counts of viable cells. Reproduced from Castelletto et al. [2016].

was detected 22 ± 2 hours after transfection. Its knockdown "fitness" (Figure 1.12, b) was computed relative to cells treated with siRNA alone (background) and normalised against viable cell counts. Capzip fitness was lower than Lipofectamine® one, however cells treated with capzip remained viable longer, suggesting that capzip has little cytotoxicity. The experiment was performed at neutral to positive charge ratio close to one (where each siRNA molecule has a -42 e charge and capzip a +6 e charge). Test experiments performed at higher peptide-to-siRNA ratio showed no improved uptake.

Finally, the peptide exerted an antimicrobial function: both the L- and D- non-assembled peptides were effective against Gram-positive and negative bacteria (*S. aureus*, *B. subtilis*, *M. luteus* - positive - and *E. coli*, *P. aeruginosa*, *S. typhimurium*, *K. pneumoniae* - negative). The minimum inhibitory concentrations (MIC), the low-

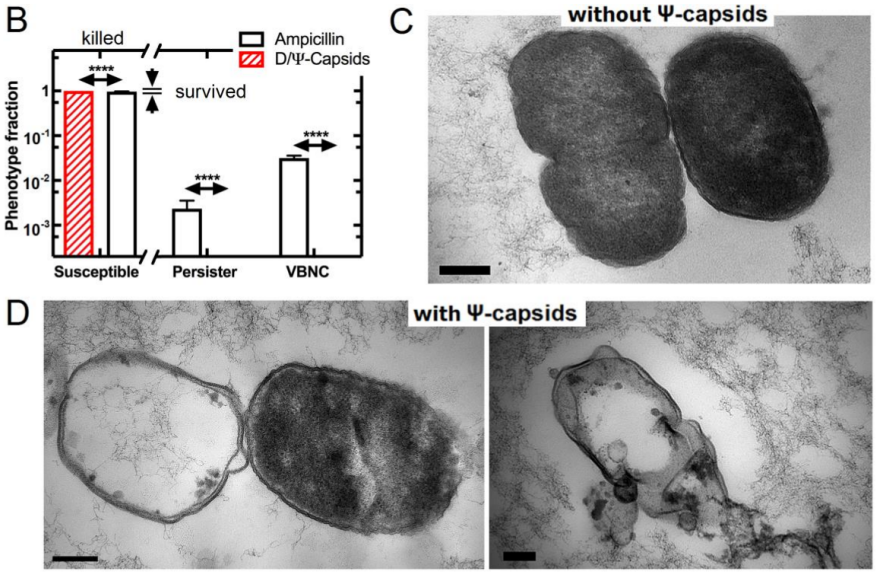


Figure 1.13: B: *E. coli* cell fractions treated with D-capsids and ampicillin (means and standard error obtained for 3332 cells hosted in 2331 independent microfluidic channels). C and D: electron micrographs of microtomed *E. coli* cells before (C) and after (D) treatment with D-capsids (scale bar 200 nm). Reproduced from Kepiro et al. [2019].

est concentration which prevents visible growth of the bacterium, was typical of other antimicrobial agents. This antimicrobial action extended also to cell types which are not susceptible to ampicillin (Figure 1.13, B). Additionally, no haemolytic effects were observed on human erythrocytes, i.e. the concentration required to achieve 50% cell death compared to untreated cells (LC_{50}) was larger than 10 times the MIC necessary to kill most bacterial species.

To better understand the type of action capzip performs on the bacteria, electron micrographs of *E. coli* cells were taken before and after the treatment. Figures 1.13 C and D show the membrane disruption and the leakage of intracellular material after the introduction of capzip. A more detailed view of the effects of a capzip capsule on the bacterial membrane was provided by experiment on Supported Lipid Bilayer with negative total charge (mixed DLPC and DLPG, 3:1 ratio). Deposition of capsules on them created localized pores within minutes, as proven by AFM experiments repeated in time. The pore depth ranges between 1.4 and 2.2 nm, which is smaller than the radii of the capsules, but is sufficient to disrupt the structure of the membrane (Figure 1.14).

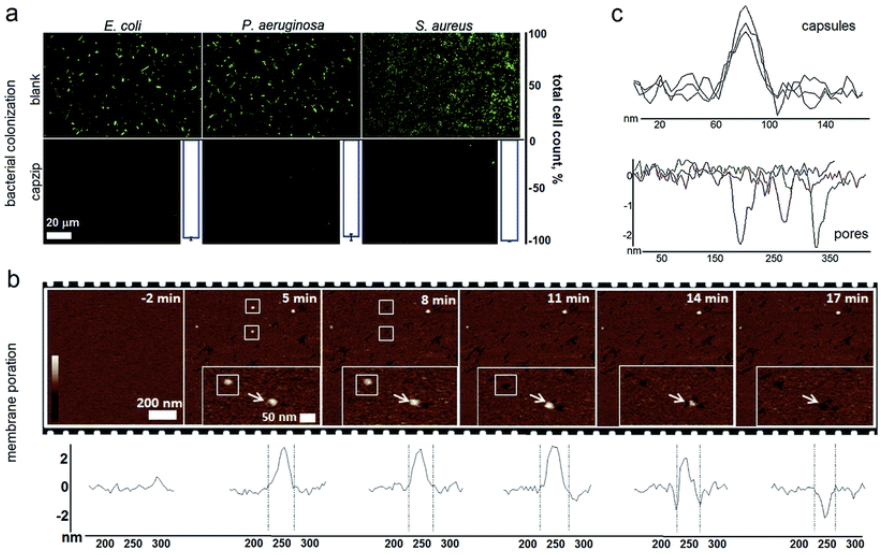


Figure 1.14: Antimicrobial activity of capzip capsules on bacterial model membranes (Supported Lipid Bilayer). (a) Confocal micrographs of stained bacterial cells after 16 hour incubations with and without capzip. White histogram bars denote total cell counts (%) for bacterial colonization with capzip after subtracting background adhesion taken as 100%. (b) AFM topography of SLBs during capzip incubation in solution. White boxes and arrows highlight individual capsule conversions into pores (color scale 6 nm). Cross sections show the evolution of this capsule into a pore in real time. (c) Representative cross-sections of capsules and pores. Reproduced from Castelletto et al. [2016].

Finally, to prove the viability of capzip as antimicrobial agent in vivo, it was used to counteract methicillin-resistant *S. aureus* (MRSA) infections in G. mellonella larvae. The particular bacterial strain used was susceptible to vancomycin, which could be used as control: the larvae treated with capzip showed survival rates significantly higher than the untreated control, and comparable to those treated with high doses of vancomycin [Kepiro et al., 2019] (see Appendix A.1).

Open questions Despite the success of the experiments mentioned above, there is much information still to be uncovered on the precise action of such peptide. Specifically, both the assembly process and the antimicrobial mechanisms contain some unknown.

Regarding the former, it is important to understand (a) which amino acids or substructures allow the pairing of molecules, (b) whether such pairing is specific or not, (c) how reversible it is, and (d) how rigid the final structure is. Regarding the latter, it must be highlighted (a) what moieties in the membrane the peptide binds to, and

(b) how this binding affects the full membrane structure.

Finally, as there is evidence that the assembled molecule is a more powerful antimicrobial compound than the single one, it is interesting to understand whether any cooperative action is taking place, or the enhanced antimicrobial power of the assembly is due only to the localised higher concentration of antimicrobial sequences.

Even if further experiments or future improvements in the techniques already employed might tackle some of the aspects above in the future, arguably no experimental outcome can provide an atom-by-atom knowledge of the processes of interest in any time soon. Ideally though, one would like to track with the finest level of details, both in space and time, the processes happening in any of the environments capzip has been exposed to (physiological solution, supported lipid bilayers, bacterial extracellular matrix, mammal cell membrane and cytoplasm). The impossibility of pursuing that leaves large gaps in the understanding of the system.

1.6 A computational approach to understand capzip

The gaps mentioned in the characterisation of capzip prompts for new investigations in order to complement the knowledge already provided.

Beside the quest to enrich the fundamental knowledge on self-assembling peptides and antimicrobial ones, the understanding of this very system is crucial for its further development. We already outlined in Section 1.2.4 how AMP design can proceed from already viable templates and empirical principles, when first principles are not available. Similar rules hold for designing self-assembling peptidic materials, to obtained tailored delivery vehicles (see Section 1.4.1). Therefore, a full knowledge of the interactions between capzip molecules and between their assembled structures and the membrane will drive the engineering of new likewise peptides. A knowledge-driven design would hopefully provide new blocks suitable for a double action as the one capzip performs, and this in a shorter amount of time than a research based on a less informed trial-and-error procedure of mutations in the chemical composition of the molecule. A few examples of possible knowledge-related improvements include the following:

- understanding the molecule-molecule interactions classifies the robustness of the

assembled structure and the possibility of designing blocks which disrupt under particular chemical conditions only;

- the knowledge of capzip binding mode to the bacterial membrane might suggest its suitability as a broad range spectrum compound versus the possibility of tuning its action against specific pathogens;

- querying the electrostatic profile of the assembled structure suggests which type of molecules, other than siRNA, could be efficiently co-assembled and thus delivered.

In recent years, computational techniques are becoming increasingly accessible and sophisticated and can usefully complement incomplete experimental knowledge to give more insights into how biological systems work (see Chapter 2). For this reason, it seems natural to employ such techniques to study the capzip system as well. Zooming into the details of the interactions can be performed via a theoretical modelling of the system, and thus through the simulation of its evolution in time, starting from the knowledge of the chemical composition of its parts. The technique this work focuses on is Molecular Dynamics simulations, which aims at reproducing the behaviour of a system of atoms in a classical description using basic physical laws, as it will be described in detail in the next chapter.

Thus, it is the aim of this thesis to prove that Molecular Dynamics simulations can clarify the assembly mechanisms of capzip and its interactions with biological membranes, in order to gather more information on the system and contribute in the future to the design of new molecules with enhanced functional capabilities.

Chapter 2

Methods

Molecular Dynamics simulations have been rightly defined as the 'Computational Microscope' [Lee et al., 2009; Dror et al., 2012] as they offer otherwise inaccessible insights into the molecular details underlying conformational changes of proteins and nucleic acids. Computational methods and tools based on MD are routinely applied in structural biology to quantitatively characterise the dynamics and thermodynamics of proteins and their complexes. MD simulations and the associated force fields are commonly used in the process of structure determination from NMR data or in theoretical structure prediction from homology models [Vogel and Huster, 2017; Heo and Feig, 2018]. In particular, simulations of the structure help in relieving the artefacts deriving from the experiments, or in determining the correct conformation in cases when the experimental measure or the homology modelling has a large uncertainty.

Modelling and simulating a biological system consists in describing its components and their mutual interactions, by implementing the laws of physics in the attempt to reproduce the dynamics of phenomena observed in nature. A quantum mechanics description would be the most accurate, but it is computationally too expensive to achieve for large systems. To facilitate the task, several simplified models have been devised, each suitable to investigate particular cases. In particular, a popular approach is a classical mechanics description of the dynamics (classical : quantum mechanics effects are relevant on very short time and length scales only. Therefore, for systems of biological relevance (at the microsecond time scale and nanometer length scale) the classical approximation is sufficiently accurate, and a model computationally affordable.

Another computational method for tackling similar problems is to adopt a probabilistic approach rather than MD: instead of using the laws of mechanics to evolve the system, a random sampling of its states is performed to find the most energetically favoured ones. This procedure, the Monte Carlo (MC) method, partially relieves the computational burden with respect to a dynamic approach, as it does not need to compute the forces acting on the particles, but only the energy of the system. MC simulations have been successfully used in situations where the system can get trapped in an energy minimum if simulated through MD, while it is deemed that many more minima are present, and thus a more extensive sampling is sought [Liu et al., 2012]. However, MC does not capture dynamical properties and thus can be unsuitable for some questions posed on biological systems. Despite the focus of the present work is Molecular Dynamics, it is important to be aware of alternative computational techniques as a complementary tool which can boost the understanding of the system.

We briefly present here the core theory and implementation of classical MD simulations, especially discussing the parametrisation of the different force fields used in this work, as a consistent part of it focuses on testing and implementing MD simulations parameters. Understanding their methodology provides an interpretative key with which simulations must be designed, run and interpreted in each specific case [van Gunsteren et al., 2006]. We will also discuss the strengths and limitations of MD simulations and their comparison with experiments. A review of relevant successes of MD simulations will complete the chapter.

2.1 Algorithms for Molecular Dynamics: the Newton's law

In a classical MD framework, Newton's second law of motion rules the dynamics, stating that the acceleration $\mathbf{a}$ that a particle is subject to at time t, depends on the total force $\mathbf{F}$ acting on the particle itself and on its mass m (bold denotes vectorial quantities):

$$\mathbf{F}(t) = m \cdot \mathbf{a}(t)\,. \tag{2.1}$$

As the acceleration $\mathbf{a}(t)$ is the second derivative of the position $\mathbf{r}(t)$ with respect to time, given the initial position and velocity of the particle ($\mathbf{r}(t_0)$, $\mathbf{v}(t_0)$), their temporal

evolution can be computed integrating $\mathbf{a}(t) = \mathbf{F}(t)/m$ as follow:

$$\mathbf{v}(t) = \mathbf{v}(t_0) + \int_{t_0}^{t} \frac{\mathbf{F}(t')}{m}\, dt' \,; \tag{2.2}$$

$$\mathbf{r}(t) = \mathbf{r}(t_0) + \int_{t_0}^{t} \mathbf{v}(t')\, dt' + \int_{t_0}^{t} \int_{t_0}^{t'} \frac{\mathbf{F}(t')}{m}\, dt''\, dt' \,. \tag{2.3}$$

In the case of complex biomolecular systems with many particles and multiple interactions acting between them, it is impossible to integrate analytically Equation 2.2, while a different and feasible approach consists in discretising it. The idea is to consider very short time steps of length Δt, so that in such intervals the forces are (almost) constant, and the integration of Equation 2.2 becomes trivial. A careful choice of the values to integrate allows to reduce the approximations derived from such approach. For example, choosing the velocity value at time $t_0 + \Delta t/2$ (and not at t_0) decreases the error down to orders of $(\Delta t)^4$ (rather than $(\Delta t)^2$). This framework is at the basis of the so-called leap-frog algorithm, which is used in the vast majority of MD engines:

$$\mathbf{v}\left(t_0 + \frac{\Delta t}{2}\right) = \mathbf{v}\left(t - \frac{\Delta t}{2}\right) + \frac{\mathbf{F}(t)}{m}\,\Delta t \,; \tag{2.4}$$

$$\mathbf{r}(t_0 + \Delta t) = \mathbf{r}(t_0) + \mathbf{v}\left(t_0 + \frac{\Delta t}{2}\right)\,\Delta t \,. \tag{2.5}$$

This algorithm can thus "solve" every possible Newton equation, at the expenses of some precision.

Constraint algorithms Considering that MD deals with bonded atoms belonging to multi-atoms molecules, the length of these bonds must be kept within a physically meaningful value. As the approximate procedure mentioned above may however give raise to unphysical configurations, constraint algorithms are applied after the update of atoms positions, to restrict the change in bond length. The ones used in this work are SETTLE [Miyamoto and Kollman, 1992] and LINCS (Linear Constraint Solver) [Hess et al., 1997]: the first is an exact implementation of the solution for rigid bodies of three elements only (such as water molecules in their atomistic description), while the second finds an approximate solution for molecules with more atoms.

Thermostats and barostats In addition to the constraint algorithms listed above, specific algorithms are needed to realistically reproduce the simulation's conditions of choice in terms of Temperature and Pressure.

In most cases, we are interested in simulating a system at constant temperature, which corresponds to a constant average kinetic energy E_k (with the average performed over time). Despite this condition seems less restrictive than imposing a constant E_k at each time, the fluctuations from the average value must be distributed according to a specific statistics.

Statistical mechanics has investigated the problem identifying the possible states of a closed mechanical system in thermal equilibrium with a heat bath at a fixed temperature [Gibbs, 2010]. This states form the so called *canonical ensemble*, which can be completely defined knowing the number N (and identity) of the particles, the temperature T, and the volume V of the system itself. Many other *ensembles* are possible, according to the restrictions imposed on the system: for example, if the volume V is allowed to vary but it is kept constant the pressure P (a quantity we will define in the following paragraphs), we obtain the *isothermal-isobaric ensemble*, defined by N, P and T. Again, a constant number N, volume V and total energy E define the *microcanonical ensemble*. These are the most common situations one wishes to reproduce in a Molecular Dynamics simulations, and thus effort has been directed to ensure their behaviour is reproduced at best in the simulations outcome.

Regarding the situations at constant temperature, the approximations performed by the MD algorithm lead the kinetic energy to drift from its initial value. To ensure that temperature is maintained throughout the simulation, thermostat algorithms have been devised to rescale the velocities of selected particles.

The thermostats used in this work are the Berendsen [Berendsen et al., 1984] and velocity rescale [Bussi et al., 2007] ones. They both rescale the velocities of all the particles in the simulation at each time step. For Berendsen, the rescaling factor λ is computed imposing that the corrected kinetic energy E_k^* is equal to:

$$E_k^* = \frac{1}{2}(\lambda v)^2 = \left(1 - \frac{\Delta t}{\tau_T}\right) E_k + \frac{\Delta t}{\tau_T} E_k^0 \tag{2.6}$$

where E_k^0 is the target kinetic energy and τ_T is a time interval multiple of Δt which regulates the strength of the coupling. A value of τ_T equal to Δt would suppress all thermal fluctuations, while in real systems they are still allowed. Even with larger τ_T though, this thermostat gives incorrect results as it can not maintain the correct distribution of velocities during the evolution of the dynamics, thus violating the statistics of the *canonical ensemble*. To solve this problem, the velocity rescale thermostat was introduced [Bussi et al., 2007]. It adds to the right side of Equation 2.6 a stochastic term:

$$2\sqrt{\frac{E_k\, E_k^0}{\tau_T\, N_f}}\, dW \tag{2.7}$$

with N_f the number of degrees of freedom in the system and dW a Wiener noise [Durrett, 2010] which ensure fluctuations are sampled correctly.

The other relevant quantity, pressure, is defined as the average quantity of motion exchanged between the particles and the walls of the box they are confined to, which depends on their frequency of collision. Most MD simulations are run under periodic boundary conditions, i.e. a particle which exits from the simulation box during a move is brought back on the opposite side of the box. This mimics the presence of an infinite number of equivalent boxes one next to the other, and alleviates the finite size effects implicitly modelling an infinite simulation box. As particles are not effectively bouncing on the walls, the pressure is computed from the velocities of the ones trespassing the box boundaries during a move. Barostat algorithms rescale the size of the simulation box to adjust the collision frequency and thus the pressure to the target value.

As a note, pressure is an intensive quantity, i.e. a local physical property of a system and does not depend on the system size or the number of particles in it (extensive quantities). However, both these quantities are involved in its computation in the framework used. From the classic kinetic theory of gases, for a homogeneous gas of N molecules of mass m in a box of volume V, in isotropic conditions, the pressure can be computed as:

$$P = \frac{Nm\langle v^2\rangle}{3V}. \tag{2.8}$$

Thus, rescaling the size of the system (extensive) and dividing N (extensive) by this quantity, will influence the pressure (intensive).

Usually all the box dimensions are rescaled by the same amount. In the case of anisotropic systems like lipid simulations, an anisotropic coupling can be enforced: for these systems, the membrane patch is usually parallel to one side of the simulation box, and spans the whole side area. In this way, it joins with its periodic images, without any water molecules in between. This is convenient as it shields the hydrophobic core of the membrane from water, moreover it ensure that the membrane remains planar. Indeed, isolated patches of membrane tend to fold into vesicles to minimise their energy. In the situation where the membrane joins its periodic images, to compute pressure, the box sides parallel to the membrane plane are rescaled separately with respect to the one perpendicular to it. This is because lateral pressure exerted by the lipids is expected to be different with respect to the one of water.

The barostats used in this work are the Berendsen [Berendsen et al., 1984] and Parrinello-Rahman [Parrinello and Rahman, 1981] ones. The Berendsen barostat is analogous to its thermostat counterpart, as it defines a scaling factor for the velocities (and thus the coordinates) based on the target and effective pressure P_0 and P, and a coupling time τ_P:

$$v^* = v\left(1 - \frac{\beta\,\Delta t\,(P_0 - P}{\tau_P}\right)^{1/3} \tag{2.9}$$

where β is the theoretical isothermal compressibility. As also the Berendsen barostat suffers from problems regarding the correct sampling of fluctuations, the Parrinello-Rahman is used for production runs. It introduces the constraint of constant pressure in the dynamical equations of the systems (in a Lagrangian approach) which allows to reproduce correctly the *isobaric ensemble*.

It must be noticed that the use of the correct barostat is crucial when measuring, e.g., the true isothermal compressibility of the system K (as opposed to the theoretical value β introduced in Formula 2.9). This quantity is defined as the relative volume change as a response to a pressure change, and is directly related to the fluctuations of the volume at constant pressure:

$$K_T = \frac{\langle \Delta V \rangle^2}{V k_B T}. \tag{2.10}$$

Clearly, an incorrect distribution of fluctuations would produce a wrong estimate of

this quantity.

Finally, when choosing a thermostat and barostat to be used in conjunction, one must check in the literature for their mutual compatibility to obtain the correct *isothermal-isobaric ensemble*. For this reasons indeed, some pairings are not implemented in the standard MD engines.

2.2 Force fields

Force fields for classical MD simulations provide the expression of the potential energy of a system. Thus they determine the forces employed in Newton's law ruling the dynamics. They usually rely on the breakdown of interactions into several independent, additive and derivable terms, identified on an empirical physical basis.

We report here the functional form of the GROMOS force field [Oostenbrink et al., 2004; Schmid et al., 2011; Reif et al., 2012] implemented in the GROMACS MD simulation engine [Berendsen et al., 1995; Abraham et al., 2015, 2018], as an example of a classical force field. Other force fields can have slightly different implementations, however the general classification of interactions and the type of functional forms used are similar.

Covalent (bonded) interactions Covalent interactions are modelled with potential energy terms representing bond stretching, angle bending, improper and proper dihedral angles torsion. The functional form of the potential energy function for each of them aims at a simplified, classical description of the atomic motion of molecules. Often, it is modelled as a harmonic-like vibration around the equilibrium position, regulated by a constant. In the GROMOS force field, this translates in the equations displayed in Table 2.1, where for proper dihedrals, the convention states that ϕ_{ijkl} is the angle between the (i, j, k) and (j, k, l) planes; with i, j, k, and l four subsequent atoms, for example along a protein backbone. A value of zero for ϕ_{ijkl} corresponds to a *cis* configuration (penalised) and π to a *trans* (favoured). The integer n denotes the number of equally spaced energy minima available in a 360° turn. The same conventions hold for improper dihedrals ξ_{ijkl}, which are used to ensure ring planarity and control the chirality of tetrahedric centres.

Type	**Eq. pos.**		**Const.**	**Functional form**
Bond	b_{ij}	k^b_{ij}	$\frac{\text{kJ}}{\text{mol m}^2}$	$V_b(\mathbf{r}_{ij}) = \frac{1}{4}\, k^b_{ij} \left(\lvert\mathbf{r}_{ij}\rvert^2 - b^2_{ij}\right)^2$
Angle	θ^0_{ijk}	k^θ_{ijk}	$\frac{\text{kJ}}{\text{mol}}$	$V_a(\theta_{ijk}) = \frac{1}{2}\, k^\theta_{ijk} \left(\cos\left(\theta_{ijk}\right) - cos\left(\theta^0_{ijk}\right)\right)^2$
Dihedral	ϕ^0_{ijkl}	k^ϕ_{ijkl}	$\frac{\text{kJ}}{\text{mol rad}^2}$	$V_d(\phi_{ijkl}) = k^\phi_{ijkl} \left(1 + \cos\left(n\,\phi_{ijkl} - \phi^0_{ijkl}\right)\right)$
Improper	ξ^0_{ijkl}	k^ξ_{ijkl}	$\frac{\text{kJ}}{\text{mol}}$	$V_{id}(\xi_{ijkl}) = \frac{1}{2}\, k^\xi_{ijkl} \left(\xi_{ijkl} - \xi^0_{ijkl}\right)^2$

Table 2.1: Functional form of bonded interaction in the GROMOS force field implemented in GROMACS.

It must be noticed that these types of potentials cannot model the rupture of a bond.

Non bonded interactions Non bonded interactions include the short range Pauli repulsion, the van der Waals attraction, and the long range electrostatic term.

The first two can be modelled together by a Lennard-Jones potential. Its functional form, describing the interaction between two neutral atoms at distance r, models the long range dispersion with a r^{-6} behaviour typical of the dipole-dipole interactions found in noble gases (London dispersion forces), while the Pauli term is approximated by a r^{-12} behaviour to ease the computation in relation with the previous one:

$$V_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]. \tag{2.11}$$

Two parameters, ϵ and σ, tune the interaction strength and the equilibrium distance. They are fitted against experimental data and are specific of each pair of atoms species.

The Coulomb energy between two charges q_1 and q_2 at distance r is represented by the Coulomb law:

$$V_C(r) = \frac{1}{4\pi\epsilon_0}\,\frac{q_i q_j}{\epsilon_r r_{ij}} \tag{2.12}$$

with ϵ_0 the dielectric constant of vacuum and ϵ_r the relative dielectric constant, introduced to properly take into account the screening provided by the material surrounding the object (usually water), as polarisability is not included in this description.

The treatment of non-bonded interactions requires particular care because of their long range nature: in every point of the simulation box many forces from distant atoms

are acting at the same time, making the prediction of the outcome difficult. The van der Waals forces decay fast, therefore the tail of their functional can be cut after a threshold distance with little impact on the outcome; while Coulomb interactions, with their slower decay, must be taken into account throughout the whole simulation box. Many algorithms have been devised to efficiently compute them, like the Particle Mesh Ewald [Essmann et al., 1995] or the Reaction Field [Tironi et al., 1995] approaches.

Finally, all biomolecular force fields, and in particular their van der Waals interactions, are parametrised to describe systems at room temperature, therefore simulations performed at substantially different temperatures must be interpreted carefully.

2.2.1 Classification of force fields

Many force fields for classical MD simulations adopt a functional form equal or similar to the one described above. Their difference lies in the number of degrees of freedom modelled, in a hierarchy of descriptions proceeding from detailed to coarse (Figure 2.1). Three possible classes of descriptions are:

- all-atoms force fields, where all the atoms are present in the description, and represented as spheres of variable size according to their van der Waals radius (e.g. proportional to σ in a Lennard-Jones model). Examples of all-atoms force fields are AMBER [Maier et al., 2015; Dickson et al., 2014; Wang et al., 2004], CHARMM [MacKerell et al., 1998; Klauda et al., 2010; Huang and MacKerell, 2013] and OPLS all-atom [Jorgensen and Tirado-Rives, 1988].

- united-atoms force fields, similar to the previous ones but where non-polar hydrogens are incorporated in the heavy atom they are bonded to. The "united" atom is given a new σ parameter and increased mass according to how many hydrogen it includes. The GROMOS force field [Oostenbrink et al., 2004; Schmid et al., 2011] follows this philosophy, and OPLS has also a united atom version [Jorgensen et al., 1996].

- coarse-grained force fields, which group together in one unique bead few atoms, to reduce the number of variables to compute. The clustered atoms are such that their mutual distances are expected to vary little with respect to the ample

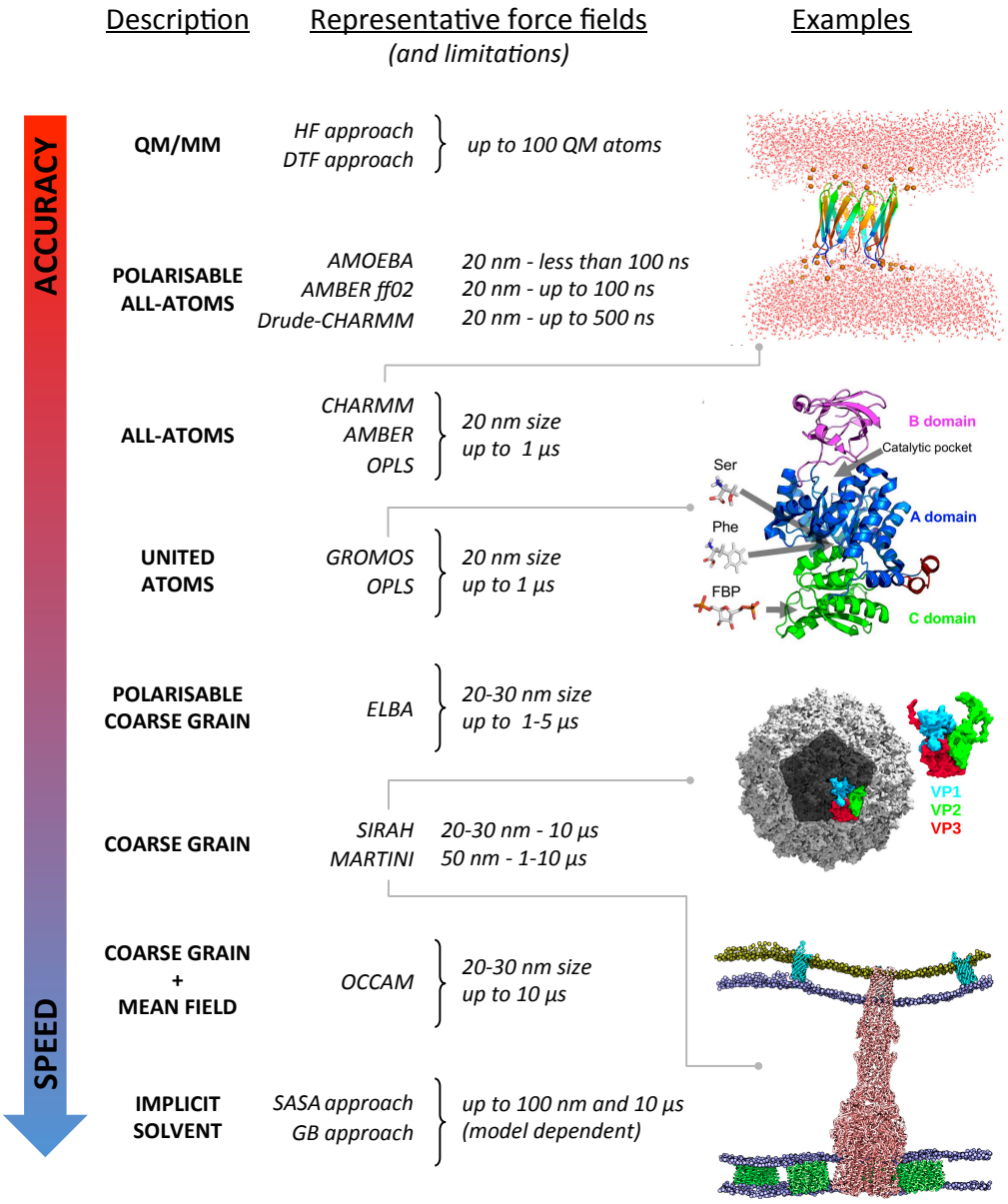


Figure 2.1: List of popular simulation force field for biomolecules, ordered from detailed to coarse (reference to the relative papers in Section 2.2.1). On the left, snapshot of notable systems simulated with the force fields CHARMM (adapted from Lipkin et al. [2017]); GROMOS (adapted from Macpherson et al. [2019]); SIRAH (adapted from Machado et al. [2017]) and MARTINI (adapted with permission from Samsudin et al. [2017]).

movements of components of the system far away from each other (which will be grouped in different beads). The MARTINI [Marrink et al., 2007; Monticelli et al., 2008; de Jong et al., 2013] and SIRAH [Machado et al., 2019; Barrera et al., 2019] force fields belong to this category.

We now give a more detailed insight of the characteristics and parametrisation strategies of the atomistic and the coarse-grained force fields employed in this work.

2.2.2 The GROMOS force field

All-atom and united atom force fields are parametrised against first-principle or experimental values. While for the all-atom force fields AMBER and CHARMM the parametrisation is based on *ab initio* quantum mechanics calculations refined against experimental data [Maier et al., 2015; Dickson et al., 2014; Wang et al., 2004; MacKerell et al., 1998; Klauda et al., 2010], the united-atom GROMOS force field relies on the reproduction of heat of vaporization of small molecules and free enthalpies of solvation of small compounds in different solvents, at physiological temperatures and pressures [Oostenbrink et al., 2005; Schmid et al., 2011; Reif et al., 2013]. These procedures sets not only the constants of the bonded interactions, but also the partial charges of the atoms inside a molecule: as no electrons are included for the sake of efficiency, their redistribution across atoms which are bonded is modelled through fractional charges assigned to each atom (while the total charge of a molecule must sum to an integer). Moreover, it is assumed that the parametrisation performed for small moieties can be transferred to a larger compound including these moieties. This limits the number of chemical groups to be described in order to simulate biomolecules.

In every MD simulation, the description of water is crucial. Out of the many water model proposed, the GROMOS parametrisation has been performed with a flexible simple point charge (SPC [Berendsen et al., 1981]) model. This description represents water as a three atoms molecule, with a negative charge on the oxygen and positive complementary charges on the two hydrogen atoms, and allows flexible hydrogen-oxygen bonds. This model reproduces correctly the density and dielectric permittivity of water [Mark and Nilsson, 2001]. To be noticed that, from the point of view of the compu-

tation, water molecules are the vast majority of the particles involved in a simulation and thus a significant fraction of the computer time is spent in updating their positions and calculating solvent-solvent interactions.

The improvement of computational techniques and reparametrisation strategies prompts the periodical release of newer versions of force fields. In the present work, we employed version 53A6 of the GROMOS force field [Oostenbrink et al., 2004] for the set of simulations involving peptidic assembly in solution, while we switched to 54A7 [Schmid et al., 2011] and 54A8 [Reif et al., 2013] for the simulations involving biological membranes. While it is advisable to always have a coherent set of parameters across simulations, to compare their outcome in a consistent manner, when extending the system simulated to include membranes, we deemed the newer parameter sets more suitable because of the improvements introduced in the phosphocholine head parametrisation [Marzuoli et al., 2019] (see Chapter 4 for a complete discussion on lipid parametrisation in GROMOS).

2.2.3 The SIRAH force field

The first coarse-grained force field we introduce groups multiple atoms in one bead but aims at maintaining chemical and structural details of the biomolecules described. As such it sets itself between the atomistic GROMOS description, and the coarse-grained MARTINI force field [Marrink et al., 2007; Monticelli et al., 2008; de Jong et al., 2013] which will be introduced in the following.

Two different approaches are taken to develop a coarse-grained force field: top-down and bottom-up. In the first, parameters are fitted directly to global quantities derived from experiments, as performed in the atomistic GROMOS parametrisation. In the second, coarse-grained simulations results are fitted to outcomes from atomistic ones.

SIRAH [Darr et al., 2015; Machado et al., 2019; Barrera et al., 2019] is a top-down force field derived to fit structural properties of proteins. It aims at reducing the complexity of an atomistic description while still being able to reproduce the correct secondary structure of proteins across a wide variety of folds contained in the PDB, together with a correct representation of their dynamics.

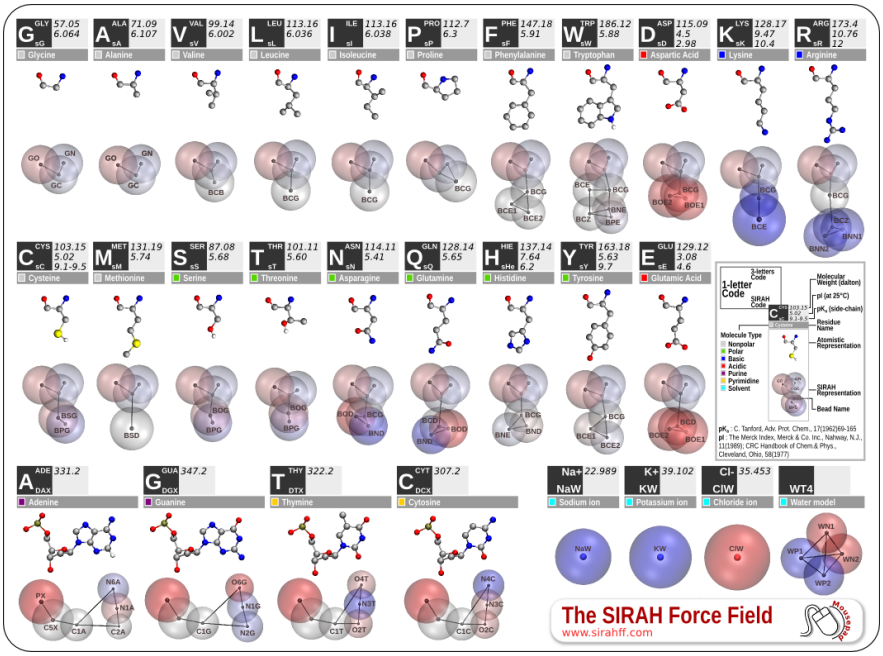


Figure 2.2: Description of amino acids and nucleic acids in the SIRAH force field. Reproduced from Institut Pasteur de Montevideo Group of BioMolecular Simulations [2019].

To obtain this, it opts for a non-uniform granularity, i.e. according to the region of interest a different number of heavy atoms is grouped in a bead, from a minimum of two up to four. In the case of proteins, it maintains the backbone flexibility by grouping NH, $C_\alpha H$ and CO in three different beads, while side chains are represented with less details, generally grouping three atoms together. A schematic of the mapping for each amino acid is shown in Figure 2.2. Contrary to force fields where the amino acid backbone is mapped to one bead only, the SIRAH description allows to reproduce secondary structures without recurring to additional constraints. The dual granularity approach is based on physico-chemical intuition, and is more difficult to generalise than a uniform one. Nevertheless, the force field has been recently extended to lipids [Barrera et al., 2019], while it comprised a parametrisation for DNA molecules since its infancy.

The modelling of water in a coarse-grained force field is also critical: usually, a few water molecules are grouped together in one bead. This has two implications: water particles are large and thus cannot solvate very narrow pockets; moreover, collapsing the molecules in one single point in space removes the separation of charges, and the characteristic dipole every water molecule should have is lost. The dipole of water is

responsible for hydrogen bonds formation and for the electrostatic screening observed in an aqueous solution. Such screening can be roughly modelled tuning the relative dielectric constant, but as this is a mean field approach, it cannot account for local effects. To partially obviate to that, SIRAH force field maps four waters to a tetrahedral molecule, with one bead on each vertex: all the bonds are rigid, and the structure serves the purpose of having a repartition of plus and minus charges, by assigning a positive charge to two vertices and the opposite charge to the other two, giving a polarisable structure. The geometrical arrangement reproduces the tetrahedral network of water molecules observed in its liquid state, which is characteristic of this fluid and tunes its remarkable properties (as the high specific heat capacity, and the non-linear relationship between density and temperature in its liquid state). Finally, ions are represented as spheres including the ion and the first hydration shell comprising 6 water molecules.

Based on the above premises, SIRAH force field simulations of different peptides and proteins in solution proved to match the relative NMR results, showing a good reproduction of secondary structures; simulations of lipids randomly oriented in water showed the formation of an organised bilayer, and the expected behaviour of a few transmembrane proteins in model membranes was correctly reproduced [Machado et al., 2019; Barrera et al., 2019].

2.2.4 The MARTINI force field

The MARTINI force field is a popular coarse-grained description of biological molecules [Marrink et al., 2007; Monticelli et al., 2008; de Jong et al., 2013]: developed originally with a focus on lipids, it has then been extended to include proteins, small ligands and DNA/RNA molecules 2.3.

MARTINI opts for a four-to-one approach, i.e. four heavy atoms are grouped in one bead. The number of bead types has been kept to the minimum necessary to represent biological molecules. They are organised systematically in polar, non-polar, apolar, or charged, and each type has a number of subtypes with increasing polarity to differentiate the chemical nature of the underlying atomistic structures. This systematic approach can be easily transferred to new compounds, without the need of introducing new bead types. The only exception is represented by rings molecules, where a two-to-

one approach is needed to maintain the circular topology.

The MARTINI force field chooses a top-down approach to parametrise non-bonded interactions, tuning them against experimental partitioning free energies between polar and apolar phases, while bonded interactions are derived from reference all-atom information, in a bottom-up approach.

The four-to-one mapping implies that the amino acid backbone is represented by one bead only, preventing the description of directional bonds which are key to reproduce the secondary structure (Figure 2.3(a)). The bonded parameters partially account for this, favouring for each residue type the backbone conformation in which it is most likely found (as computed from the Protein Data Bank - PDB [Berman et al., 2000]). When this is not sufficient, the protein can be constrained around a given structure through an elastic network model approach which constraints residues to keep their mutual position through harmonic potential applied on their backbone beads (ElNeDyn [Periole et al., 2009]). However, both the backbone parametrisation and the use of ElNeDyn imply that local energy minima of the natural structure are not well sampled in MARTINI simulations, biasing the understanding of the structure dynamics.

The MARTINI force field provides two water models (Figure 2.3(c)). The standard one groups four water molecules in one bead only, loosing the polarisability typical of water, the effect of which is partially restored with the use of a high dielectric constant. The polarisable water model [Yesylevskyy et al., 2010] maps instead four water molecule to a single "inflated" water, i.e. a three-beads molecule with the same shape of a single molecule, but expanded, and a charge splitting which can account for the water dipole. Overall, the MARTINI force field pushes the limits of simplification to enhance the simulations speed-up, with considerable gain in efficiency with respect to atomistic or even SIRAH simulations. Despite it can not capture some fine details of the systems studied, it has been successfully applied to describe the behaviour of many biological membranes [Khalid et al., 2019; Samsudin et al., 2017], lipid self-assembly [Marrink et al., 2007] and peptide-membrane binding [Song et al., 2019]. The (re)introduction of a more detailed water model allowed the description of electroporation processes and the translocation of ions through bilayers [Yesylevskyy et al., 2010].

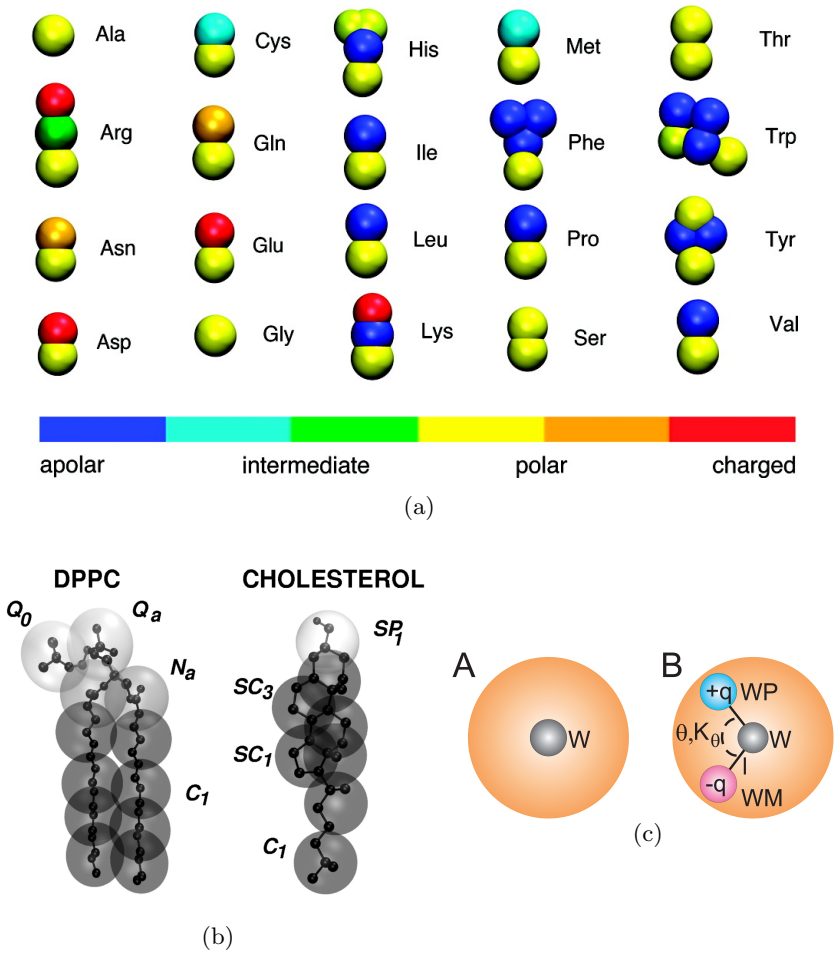


Figure 2.3: Description of amino acids (a), example lipids (b) and water models (c) in the MARTINI force field. In (c) the orange sphere represents the van der Waals radius of the central atom. Reproduced from Monticelli et al. [2008]; University of Calgary Biocomputing Group [2016]; Yesylevskyy et al. [2010]

2.2.5 Backmapping techniques

Coarse-grained descriptions are very effective in reproducing long time scales; however, to retrieve finer details after such extensive exploration, backmapping techniques have been designed to obtain atomistic configurations from the coarse-grained ones [Wassenaar et al., 2015]. These backmapped structures can in turn be simulated at the atomistic level to explore the short time scale movements around such interesting conformation. The easy conversion between the two resolution, gave rise to many multiscale studies applied to biomolecular systems [Lee et al., 2012].

2.3 Validation and challenges of MD simulations

Validation of MD simulations is performed by comparison with experiments: the same properties obtained experimentally are computed from the MD trajectory as well, and the two compared. If these are correctly reproduced, it is usually assumed that the simulation is sampling the correct ensemble of states and then one can identify in the simulation the determinants responsible for the experimental outcome of interest, which are not accessible by the experiment itself.

The comparison however is not always easy: often the experimentally measured quantity is an average in time and/or space (for example Circular Dichroism spectra or SAXS profiles of a peptide in solution) and many different combinations of computationally derived structure ensembles can produce compatible results. It is still challenging to compare experimentally derived ensembles and the ones derived by molecular simulations. Indeed the extracted average properties may be different, so it is important to understand which are the relevant ones playing a role in the measured ensemble before attempting comparisons.

Thus, in the validation of MD outcomes, it is important to have a critical attitude and to interpret the result within the validity of the approximation [van Gunsteren et al., 2008]. Indeed, agreement may arise from either a simulation that reflects correctly the experimental system; but also when the property examined is insensitive to the details of the simulated trajectory, or it can be a result of compensation of errors, which is more likely to occur for systems with a high number of degrees of freedom. Similarly, disagreement may hint at an error in the simulation (either in the model, the implementation, or simply the estimated simulation's convergence) or an error in the interpretation and/or conditions of experimental set-up (either in the result itself o its interpretation), so that both must be carefully checked to improve a convergence in the agreement. Some apparently negative results may suggest or stimulate new experimental settings to validate the hypotheses one was set to test [Gonçalves et al., 2013; Meißner et al., 2014].

Additionally, simulations still suffer from the limited computational time available: most of the times, the real experimental system is simply too large to be reproduced

and the time scale of the process too long. Simulations are thus confined to explore a restricted space, implying that the initial conditions must be chosen carefully to optimise the search and avoid any bias which might persist for the whole length of the simulation. The use of enhanced MD techniques can increase the chances of sampling relevant states [Bernardi et al., 2015]. As a non comprehensive list of these techniques, we mention replica-exchange algorithms [Best et al., 2005] which combine together multiple simulations held at different conditions; local potential-energy elevation (or metadynamics) [Barducci et al., 2010, 2011] which avoids the re-sampling of already visited conformations adding an energy penalty to them; umbrella sampling [Mills and Andricioaei, 2008] which reconstructs free-energy barriers from simulations performed at specific values of the coordinate along which the barrier exists; or finally the simple use of higher temperature to overcome energy barriers [Kirkpatrick et al., 1983].

Finally, one should keep in mind that the force fields used are far from optimal, partly because they rely on approximate functional forms, and partly because it is difficult to find experimental observables measured with the desired resolution able to discriminate between sets of parameters.

Nevertheless, it remains important to note the contribution that Molecular Dynamics simulations have played in disclosing important details behind biophysical processes and in unravelling molecular details not accessible to experiments.

2.4 MD simulations: successes

Consistently with the focus of this thesis, we will privilege examples of simulations investigating antimicrobial peptides as well as self-assembling ones, showing how computational techniques can help the design of novel molecules with improved specific characteristics.

2.4.1 Simulations of antimicrobial peptides

MD simulations of antimicrobial peptides are quite well documented since the first developments of the technique. Such peptides are a suitable system for a computational investigation as, in most of the cases, their mechanisms of action are not completely

understood from the experimental information available (see Section 1.2.2). As experiments prove that even the mutation of one single residue in short AMPs can change remarkably the antimicrobial activity of the sequence (see Section 1.2.4), it is then clear that their action is governed by subtle atomic interactions, so that MD simulations, with their atomistic resolution, can help understanding this aspect.

Systems As mentioned in the previous chapter, it has been proposed that most AMPs act through a process of attraction to the bacterial membrane, possible aggregation with other copies of the same sequence, insertion, and membrane lysis. The time scales of the overall process (at least of the order of microseconds and up to seconds) are accessible using coarse-grained techniques, but not - or rarely - atomistic ones. For this level of description instead, the different steps are usually investigated separately, based on prior hypotheses: for example, the peptide can be positioned close to the membrane surface [Wang et al., 2012b], or placed directly within the membrane core with different insertion depths to verify which configurations are the most disruptive ones [Lipkin et al., 2017]. The full insertion process will be then reconstructed from a "stepwise" knowledge combining the different states sampled. For these reasons, the choice of the conformation to simulate, i.e. the initial conditions in terms of the mutual position of peptide and membrane, is crucial, as it likely biases the simulation towards the sampling of a particular subset of configurations. Recent advances in the computational power are making possible the simulations of the full process even at the atomistic resolution for simple enough systems [Ulmschneider, 2017; Sun et al., 2015], nevertheless the "stepwise" approach is still common and the preferred one in case of complex AMP systems.

Model membranes The second important choice in the setup of a simulations of antimicrobial peptides concerns the model of the membrane to simulate. In an effort to keep complexity low, bacterial and mammal membranes are often modelled with a minimal number of lipids. Usually, models of bacterial membrane retain as only key characteristic an overall negative charge, with about 25% of anionic lipids ($-1\,e$ charge) [Lipkin et al., 2017; Wang et al., 2012b; Zhao et al., 2018a; Chen et al., 2019].

For a model mammal membrane instead, only zwitterionic lipids are employed, with the occasional inclusion of cholesterol, as it is deemed important in describing more realistically their behaviour [Lipkin et al., 2017; Wang et al., 2012b; Zhao et al., 2018a; Chen et al., 2019; Risselada and Marrink, 2008]. Because of their simplicity, these systems are used in experiments [Castelletto et al., 2016; Tang and Hong, 2009; Glukhov et al., 2005], making possible a direct comparison with simulations. Therefore, even if these simple membranes do not model accurately the structure of the cellular envelope, simulations and experiments of these systems provide a first explanation of some steps of the antimicrobial activity, with the two techniques complementing and validating each other.

Nevertheless, attempts to model cell membranes more accurately have been pursued. This can be performed at the atomistic level [Piggot et al., 2011] but the task is especially suited for a coarse-grained description, as the inclusion of all the elements of the cell membranes results in quite large systems for which atomistic computations started only in recent years to be affordable. Accordingly, coarse-grained (MARTINI) simulations have been incorporating more and more components into model membranes, describing the bacterial inner membrane, the bacterial wall, and finally the combination of the two [Khalid et al., 2019] (see Figure 2.1, bottom). These large scale, coarse-grained simulations provide information on the mechanic characteristics of the system: for example, simulation of the outer membranes of Gram-negative bacteria combined with the peptidoglycan layer (which, in bacteria, is positioned between the two membranes) elucidated how the distance between the two is variable, thanks to the presence of Braun's lipoproteins [Asmar and Collet, 2018] which act as a bridge between them, and can bring them closer by bending and tilting. On the other hand, the permeability of membranes to ions and small compounds needs to be assessed at the atomistic level, and to access informative simulation time scales, smaller and simpler systems must be chosen for the task (e.g. the inner membrane only), often together with enhanced MD techniques such as metadynamics, umbrella sampling, and replica exchange umbrella sampling. [Sun et al., 2016; Piggot et al., 2011; Carpenter et al., 2016; Pokhrel and Maibaum, 2018].

Force field comparison Finally, simulations of the peptide interaction with a model membrane are clearly determined by the parametrisation of the force field employed for protein and lipids (and by their mutual consistency). There are multiple evidence suggesting that different force fields produce very different outcomes when simulating the same system, under the same conditions [Wang et al., 2014; Bennett et al., 2016; Sandoval-Perez et al., 2017]. This is also valid for simulations of pure lipid patches (see Chapter 4), resulting in incompatible values of area per lipid, organisation of the tails and energetic profiles across the membrane, and thus has an impact in the simulations of AMPs interacting with a membrane.

For example Wang et al. [2014] observed significantly different levels of unfolding for the AMP melittin within a model membrane at the change of the force field, namely CHARMM27 and 36 (for protein and lipids respectively) [MacKerell et al., 1998; Klauda et al., 2010], OPLS all atoms (for protein) and united atoms (for lipids) [Jorgensen et al., 1996] and GROMOS 53A6 [Oostenbrink et al., 2004]; while Bennett et al. [2016] proved that the propensity of the synthetic AMP CM15 to form pores strongly depends on the force field used but also on some extent - at least at the time of the work - on the MD engine used (GROMACS compared to NAMD [Phillips et al., 2005]).

A more systematic investigation was performed by Sandoval-Perez et al. [2017], who focussed on the reproduction of membrane-protein interactions (in GROMOS 54A7 [Schmid et al., 2011], CHARMM36, Amber14SB/Slipids [Jämbeck and Lyubartsev, 2012] and Amber14SB/Lipid14 [Dickson et al., 2014]). While all of the force fields were able to reproduce the overall positioning of transmembrane proteins in the case studies tested, they showed discrepancies for amino acid side chains insertion depth and energy. Interestingly, all parametrisations gave a very broad minimum for the insertion position of Tryptophan, in line with the different roles this amino acid is supposed to assume: either an anchoring point positioned deeper in the hyrophobic region, or a partner for hydrogen bonding with the hydrophilic heads [Chan et al., 2006].

The subtle differences between parametrisations lead to the conclusion that for every particular system tested, the comparison with at least one experimentally measured quantity would be the only way to assess the simulations performance accurately.

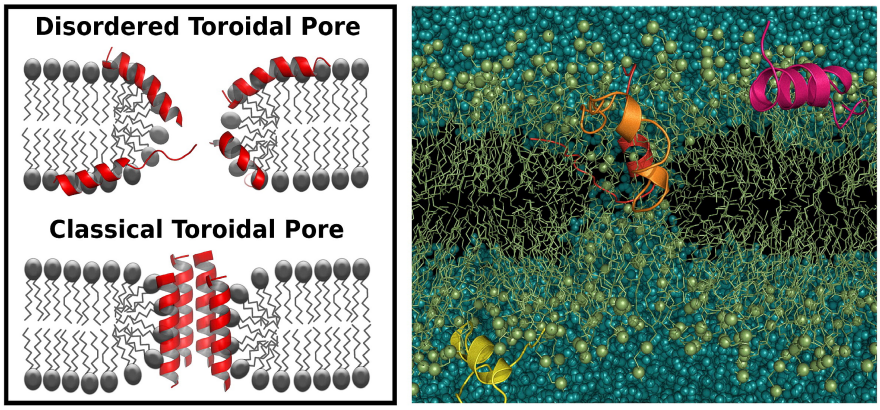


Figure 2.4: (Left) A cartoon image comparing the disordered toroidal pore state (lack of a well defined peptide orientation) to the traditional view. (Right) A snapshot of the disordered toroidal pore from simulations of melittin in DPPC. Peptides in cartoon, lipids in lines, water and ions in bead representation. Reproduced from Sengupta et al. [2008].

Simulations of membrane-peptide interaction: examples Even in the context of a simplified scenario and with the caveats coming from the parametrisation chosen, simulations of antibacterial peptides on a membrane have been successful in elucidating some AMPs mechanisms. The first important contribution consists in the introduction of the disordered toroidal pore concept: as explained in the previous chapter (Section 1.2.2), the models of membrane poration due to AMPs consist often in ordered structures where many peptides gather together to contour a pore (see Figure 1.3), and they are either in contact with the hydrophobic tails of the lipids (barrel-stave model), or with their head, as lipid molecules bend around the pore to keep their tails screened from the outside environment (toroidal pore model). However, simulations of the short helical peptide magainin MG-H2 [Leontiadou et al., 2006], among others, showed that a single copy of the helix inserted at an angle with the normal to the membrane plane, and was sufficient to displace the lipids around in a non organised manner, and form a water-filled pore (Figure 2.4).

Regarding possible rearrangements of the antimicrobial peptide structure when interacting with a membrane, simulations of cathelicidin LL-37 on pure POPG (anionic) and POPC (zwitterionic) lipid patches showed that LL-37 has a propensity to bind to the former, as expected due to the opposite charge that membrane has with respect to the cationic peptide [Zhao et al., 2018a]. However the simulations highlighted also that, in contact with POPC, the helical secondary structure was lost, while the interaction

with POPG preserved it, suggesting that the spatial arrangement of the residues, and not only the overall chemical character, is important for their action. Such type of information is hardly available to experiments or through a theoretical reasoning.

The improvements in computational resources is slowly pushing the extent of simulation time to the microsecond time scale. Thus in a recent example, the translocation of the helical PGLa peptide through the membrane has been observed as a rare event dependent on the concentration of the peptide. This event happened on the multi microsecond time scale without the formation of an organised pore [Ulmschneider, 2017]. The *in silico* experiment still benefited of an enhanced sampling in the form of a higher temperature as simulations were run at 363 K (rather than the usual 303 K), but no pre-insertion of the peptide was performed. This study shed light on a possible mechanism of permeabilisation which is usually overlooked in favour of processes involving organised channels and pores.

Similarly, simulations were able to shed light on the mechanism of translocation of Arginine-rich peptides [Sun et al., 2015]. These sequences have high positive charge, but despite this, possess a high propensity to penetrate membranes, overcoming the hydrophobic region represented by the lipid tails. Very similar peptides where the Arginines were swapped with Lysines showed no significant penetration. A commonly used explanation considers polyarginine translocation a quasi-equilibrium process, but this does not explain the selectivity against Lysines rich peptides. After extensive simulations of the two systems (multiple, hundreds of nanosecond long runs), the proposed mechanism involves the spontaneous formation of pores initiated by random (thermal) fluctuations of the position of the lipids: in some of these rare events, the transient pore would be occupied by a peptide (a precursor), which slows down the pore closure. In such situation, the translocation of other copies of the peptide is highly favoured (if their concentration is sufficiently high) as they aggregate with the precursor inside the membrane and are then pushed toward the opposite side where there is a lower charge density. Differently from polyarginines, polylysines have a much lower aggregation propensity, so that the presence of a precursor peptide inside the membrane does not induce an enhanced insertion of further peptides.

The last example brings the attention on whether oligomerisation is necessary for

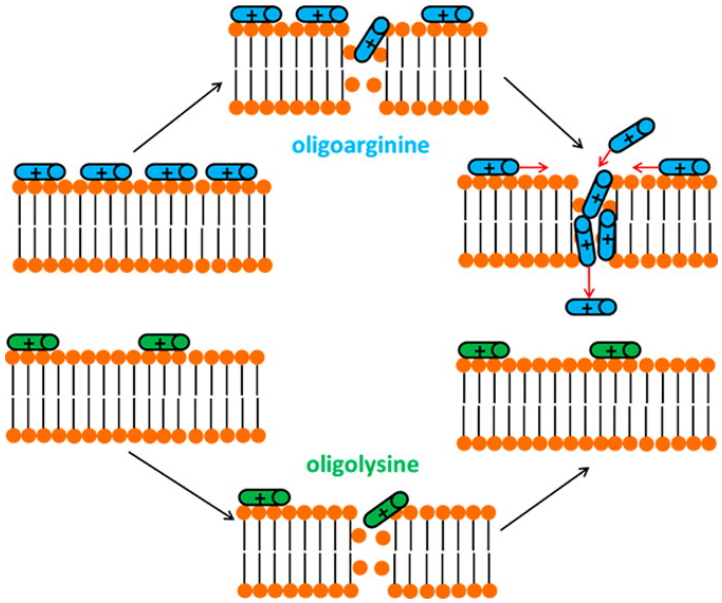


Figure 2.5: Cartoon representation of the different assembly mechanism that allow polyarginines but not polylysines to translocate through spontaneously formed pores. Reproduced from Sun et al. [2015].

an efficient antimicrobial activity. Contrary to oligomerisation in solution, which can happen on shorter time scales, the spontaneous aggregation of peptides on a membrane surface requires a long time, as the structures must diffuse on the membrane to meet each other, and many other competing processes (such as insertion) are happening at the same time. MD simulations offer insights on this aspect as well.

In a recent example, simulations of maculatin (an helical AMP) showed that the pores it forms can include a variable number of helices and thus assume many different conformations [Wang et al., 2016b]. The suggested process of pore formation proceeds via insertion of a single residue, closely followed by other ones which are able to penetrate the membrane thanks to the lipid defects already created by the first peptide.

Similar investigations were carried on also for other aggregating cell penetrating peptides, which are not antimicrobial: as such, some of them aim at inserting within cells without necessarily causing poration. One example is constituted by the influenza fusion peptides, which have been extensively studied with a simulation set up similar to the one mentioned for AMPs: a few copies of the peptide positioned on a model membrane oligomerised into aggregates of different sizes and inserted into the membrane, perturbing its local curvature [Haria et al., 2014; Collu et al., 2015].

To be noticed that, when investigating oligomerisation, the size of the system must

necessarily be increased to include all the copies necessary to form the aggregates observed experimentally. As such, with the present accessible computer time, not all the systems can be investigated from unbiased initial conditions. In the case of protegrin, a β-hairpin antimicrobial peptide which has been long though to act through the formation of transmembrane β-barrels, many variables can influence the outcome of the unknown final structure. To overcome such problems, a semi-systematic investigation has been carried on by Lipkin et al. [2017], simulating different assembly (see Figure 2.1, top). Microsecond long simulations discriminated which ones of these initial configuration formed stable pores for the whole length of the simulations, and the ones which were disrupted. As in the previous example, several different possibilities were found stable in solution, suggesting that single AMPs might have multiple mechanisms of insertion into membranes.

Most of the examples above employ atomistic descriptions of the system, but similar investigations have been carried on also using the MARTINI force field. As an example, MARTINI simulations of maculatin and aurein on POPC membranes showed different propensities for pore formation versus aggregation, showing that the coarse-grained model retains enough details and chemical information to reproduce different membrane-perturbing behaviours [Balatti et al., 2017]. Nevertheless, the developers of the MARTINI model themselves pointed out how some aspects of pore formation might not be captured in a satisfactory way [Marrink and Tieleman, 2013]. For example the penetration of water can be misrepresented, as can be expected from a model which clusters four water molecules together (indeed some pore conformations allow for the passage of fewer water molecule, if not one, at the time). As such, the MARTINI membranes can be overstabilised, and have less propensity to break and porate, with respect to their real counterpart.

In general, the outlook of simulations of antimicrobial peptides interacting with membranes goes in the direction of reproducing longer time scales thanks to the enhanced computational power available. This enhanced power would also reduce the need to use biased initial conditions or higher temperatures to speed up the simulations. Moreover, gathering the contribution of the whole community, simulations will likely go in the direction of modelling more accurately the bacterial membrane (see

previous paragraph). Finally, the force field issue must be solved in collaboration with experimentalists, finding new tests and experimental quantities to compare the computational outcome with, and make the different parameters sets converge toward a similar description of the phenomena observed, which is consistent with the experimental results.

Simulation-aided AMPs design The role of simulations in aiding AMPs design has been briefly sketched in Section 1.2.4. As pointed out, MD simulations are hardly a tool to analyse large dataset, therefore the investigation focuses on a few selected sequences.

Simulations can be helpful in integrating structural information which is otherwise lacking: such approach was followed by Liu et al. [2018] to complement the chemical features available on a dataset of short AMPs, and the overall information was used to feed a predictor of AM activity of novel sequences. Preliminary results showed that this improved significantly the ability of the predictor to discriminate whether a new sequence was suitable for antimicrobial activity or not.

Another commonly followed approach consists in using simulations to elucidate the reasons why a particular mutation is effective in terms of increased activity or decreased toxicity. Mutation screenings can be afforded experimentally on short sequences, but they are usually limited to single point mutations, as testing all the possible combined mutations is prohibitive. However, once assessed the importance of a given substitution, it is interesting to understand the mechanism behind it. For example, simulations of ovispirin and a mutant peptide with reduced toxicity showed that the bend of the helix in the latter was responsible for mitigating the interaction with mammal membranes and thus reducing haemolysis [Khandelia and Kaznessis, 2005]. In another example, temporin and a derived sequence with improved activity were investigated. Simulations showed that the mutant had a reduced aggregation propensity in water, so that more copies were ready to bind to the membrane and thus disrupt it [Farrotti et al., 2017].

In general, the protocol of integrating simulations and design is usually customised according to the system in exam, as the field has not reached yet a systematic organisation. However, it is clear that simulations used in conjunction with experimental

testing can be used to optimise already available AMP sequences, and thus contribute to device design rules for the creation of synthetic peptides with tailored properties.

2.4.2 Simulations of self-assembling peptides

Self-assembling peptides are another fascinating and challenging topic that MD simulations can help investigating. Simulating such systems implies different challenges with respect to the ones faced when simulating AMPs on membranes [Frederix et al., 2018; Orsi, 2018].

In theory, the set up of the system is quite straightforward: only the solvent characteristics and optionally the experimental salt concentration need to be matched, then choosing a random initial configuration of the molecules - in the desired concentration - would allow the simulation of the process of interest. In reality, to reproduce the dilution employed in experiments, a considerable volume needs to be simulated to host enough copies of the peptide to observe the assembly of large enough oligomers. With this approach, the time scale useful to witness a spontaneous assembly would greatly exceed the computational time available. For that reason, two main strategies have been adopted to follow assembly processes: coarse-grained simulations and/or the use of pre-assembled structures. In the following we give a few examples of these two strategies. Other routes include the choice of an implicit solvent model [Jusufi, 2013; Spaeth et al., 2011], or the use of a Monte Carlo sampling which can sometimes be less time consuming as it does not evolve the system through dynamics, but generates new configurations at random and accepts the ones energetically favoured [He et al., 2001; Majumdar et al., 2019; Luo and Robinson, 2015].

Many studies have been performed with the coarse-grained MARTINI force field to witness assembly of surfactants [Wu and Yang, 2012], polymers [Wang et al., 2012a; Bochicchio and Pavan, 2017], lipids [Lee and Pastor, 2011; Brocos et al., 2012] or peptides [Guo et al., 2012; Seo et al., 2012]. Regarding the latter, the assembly in water of peptide amphiphiles (PAs) into cylindrical fibers has been simulated [Lee et al., 2012], showing a transition from small micelles to long fibres (Figure 2.6). This example of minimal PA structures is particularly interesting for the study of AMPs as well, as it shares with them the amphiphatic character, so that having a general

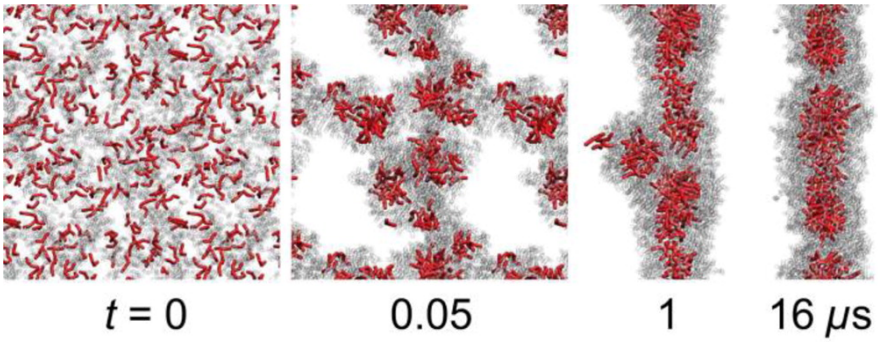


Figure 2.6: Process of peptide amphiphiles (PA) fiber formation assessed through MARTINI simulations. Hydrophobic tails in red, peptides in gray. Reproduced from Lee et al. [2012].

knowledge on how AMP analogous sequences assemble together would help in tuning their aggregation properties in water prior to the delivery to the membrane.

The second approach consists in preparing the system in pre-assemble states that are somewhat suggested by experimental evidences, and further to this using MD simulations to verify whether the conformations are kept or disrupted, and which one is the one most energetically favoured. In the work mentioned about peptide amphiphiles [Lee et al., 2012], pre-assembled fibres have been simulated at the atomistic level, to confirm the stability of the structures found with a coarse-grained approach.

The same approach has been widely employed also in cases where the final assembly was hypothesised to have a high degree of order, achievable only with a long sampling [Gudlur et al., 2012].

Similar procedures have been crucial in elucidating the assembly process of viral capsids. Capsids are very large systems and the assembly of their protein subunits is mediated by energy barriers. For these reasons, pre-assembled systems have been simulated to understand the interaction between the components and thus the first mechanisms of the assembly. This has been done recurring to ultra coarse-grained or elastic network models [Grime et al., 2016; Abi Mansour and Ortoleva, 2014] first, and only in the most recent computational advances to atomistic simulations [Perilla et al., 2016; Hadden and Perilla, 2018; Abi Mansour and Ortoleva, 2014].

The examples above show how Molecular Dynamics simulations have been employed for the investigation of many different systems, adapting the resolution, set up and

the techniques related to better query the systems of interest. Such overview suggests then that simulations would be a suitable tool to investigate the system of interest of this work, namely the self-assembling antimicrobial peptide capzip. The two aspects of its behaviour will be studied separately, adopting the necessary approximations and strategies to make the simulations efficient and to query the related questions at each time.

The details of the systems simulated and the specific parameters used for each case can be found in the relative sections of the following chapters, together with an extensive explanation of the motivation of the choices made.

Chapter 3

Capzip simulations

In Chapter 1 we introduced the molecule capzip and its properties, highlighting the unknowns of its mechanisms of action. In Chapter 2 we presented a review on Molecular Dynamics simulations, showing their past successes in elucidating the behaviour of self-assembling and antimicrobial peptides. Now, we employ this technique to understand better our system of interest. Given the complexity of the unit, and the little atomistic information available, modelling the assembly of such peptide must proceed in a stepwise manner.

The first aim is to elucidate which structures it can form in solution and what interactions are keeping the molecules together. To understand the latter it is important to retain the highest level of detail possible, and for this reason we resorted to atomistic simulations first. However, reproducing the natural assembly from a dispersed solution, as observed in the experiments, has an high computational cost when choosing an atomistic level of detail. Thus, we simulated increasingly complex pre-assembled blocks, verifying each time their behaviour in solution and inferring whether they are suitable to form a stable supramolecular structure. This approach, fully explained in Section 3.1, led to the model of a minimal capsule, which has been subsequently investigated at coarser levels to explore its behaviour on longer time scales.

This multiscale approach constitutes also an interesting methodological investigation as, to our knowledge, it compares for the first time the performances of the SIRAH force field [Machado et al., 2019; Barrera et al., 2019] versus the MARTINI one [Marrink et al., 2007; Monticelli et al., 2008], without or with polar water [Yesylevskyy et al.,

2010], at the level of protein description (while for lipids a comparison is included in Barrera et al. [2019]). Indeed, when choosing a coarse-grained representation, it is crucial to understand how the simulations results deviate from the atomistic counterpart due to the model, assuming that the atomistic is the most accurate one. This helps in selecting the appropriate description for the system and the questions that can be investigated with each coarse-grained force field.

Out of the pre-built capzip structures, a few selected ones were simulated in contact with model membranes, bacterial and mammalian, to understand the determinants of their antimicrobial activity. Details on these simulations and the specific techniques employed to promote the interaction between the peptide and the membrane are given in Section 3.2.

3.1 Modelling the assembly

Subunits simulations As previously mentioned, the antimicrobial sequence of capzip is designed with opposite charges at its extremes to favour an antiparallel β-sheets pairing with other copies of itself (Figure 3.1, A-B). MD simulations of two RRWTWE sequences paired in this fashion (Figure 3.1, B) confirm that the assembly is stabilised by the interactions between opposite charges (statistics gathered over 16 replicas, each run for 20 ns). Moreover, backbone hydrogen bonds form between Tryptophan residues of facing strands, after a shift of the mutual positions of the chains which brings Tryptophans in front of each other (see the scheme in Figure 3.2). Finally, π-stacking interaction (parallel or T-shaped) between Tryptophan side chains contributes to the interaction as well, albeit in minor measure (Figure 3.3(a)-(b); details on the computation are explained in Section 3.3). The T-shaped orientation is favoured with respect to a parallel stacking, consistently with the fact that the latter produces an unfavourable repulsion between the planar faces of the rings [Hunter et al., 2001]. Thus, Tryptophan residues dictate the interdigitation of facing residues in the two chains.

Analogous simulations of two RRWTWE sequences paired in a parallel way show loosening of the pairing and in some cases the flip of one sequence to rearrange with respect to the other in the antiparallel manner, confirming that the antiparallel ar-

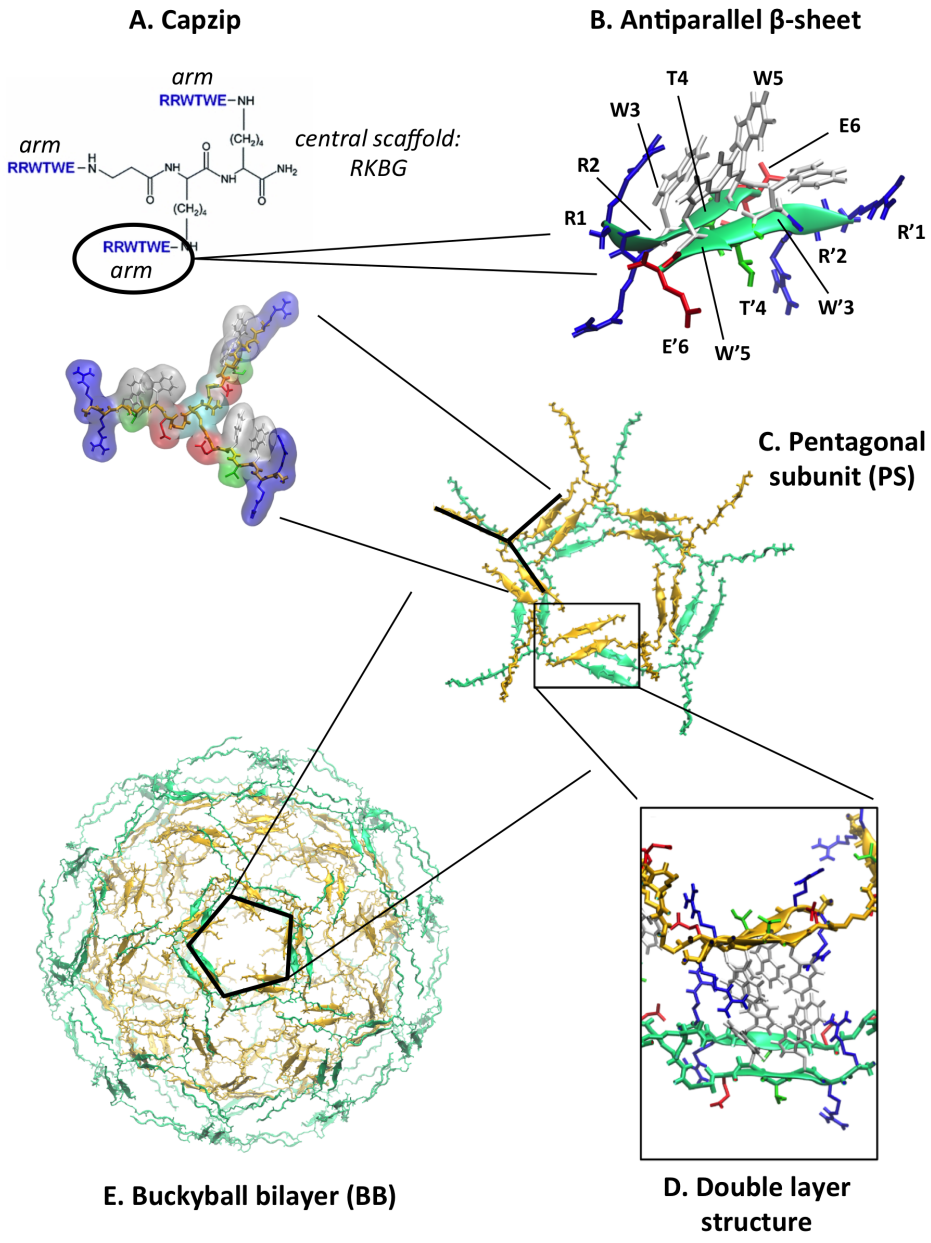


Figure 3.1: A. Capzip molecule: formula, with arms in blue characters and central scaffold (residue RKGB) in black lines; and bonds/surface representation, with surface color coded by amino acid type (blue positive, green polar, red negative, white hydrophobic). B: detail of two arms paired in an antiparallel β-sheet (bond and cartoon representation, backbone in green and side chain color coded by amino acid type). C: pentagonal subunit (bonds and cartoon representation, green and yellow for the two different layers): ten antimicrobial molecules arranged in two stacking pentagons. D: detail of the double layer of each pentagonal subunit (bonds representation of side chains is color coded by amino acid type, and backbone in cartoon representation, green and yellow for the two different layers). E: atomistic structure of the buckyball bilayer simulated (bonds and cartoon representation of backbone only, green outer layer, yellow inner layer). [VMD software, Humphrey et al. [1996]]

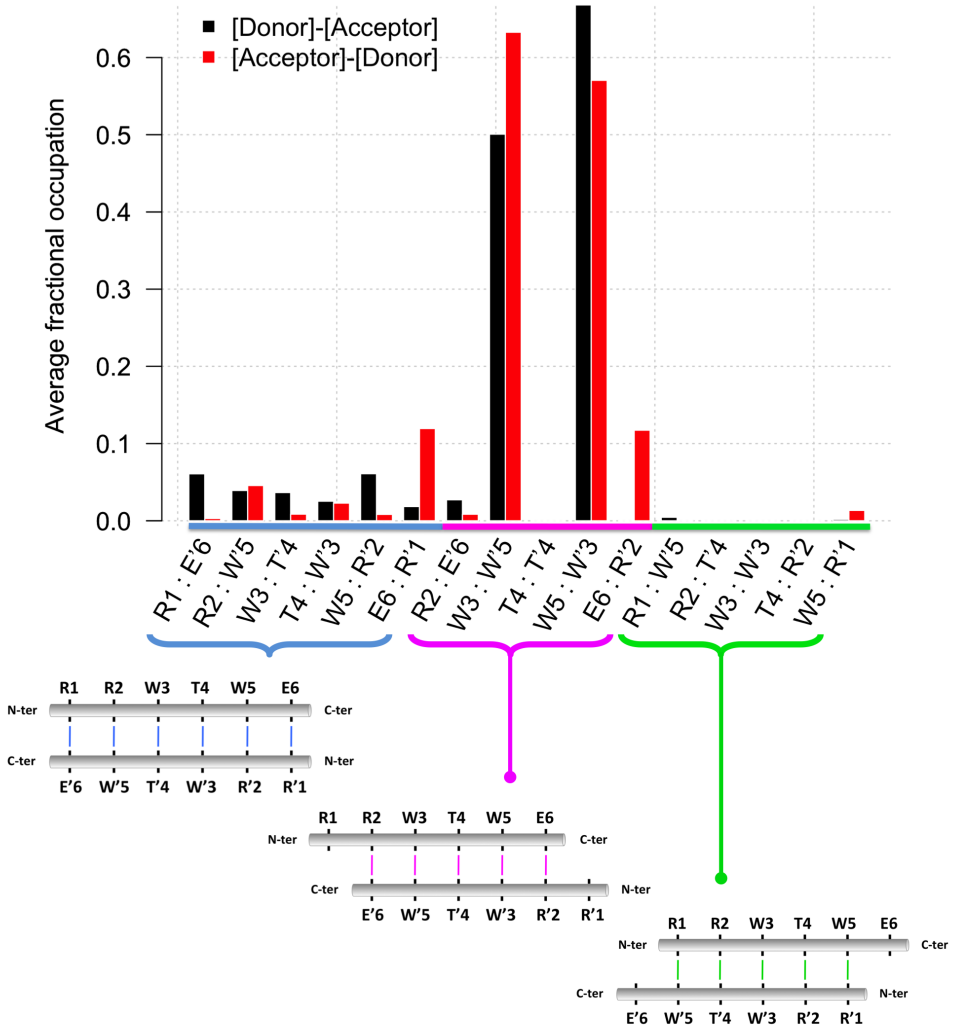
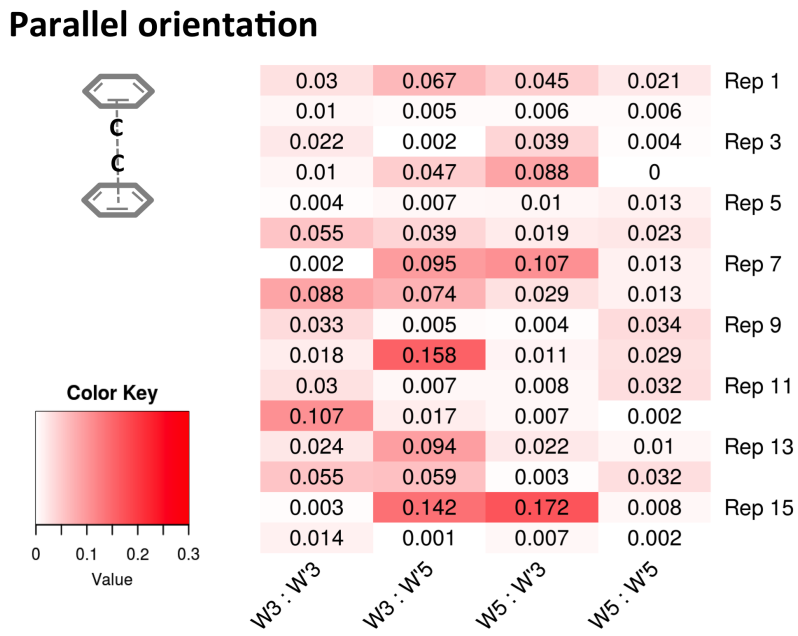
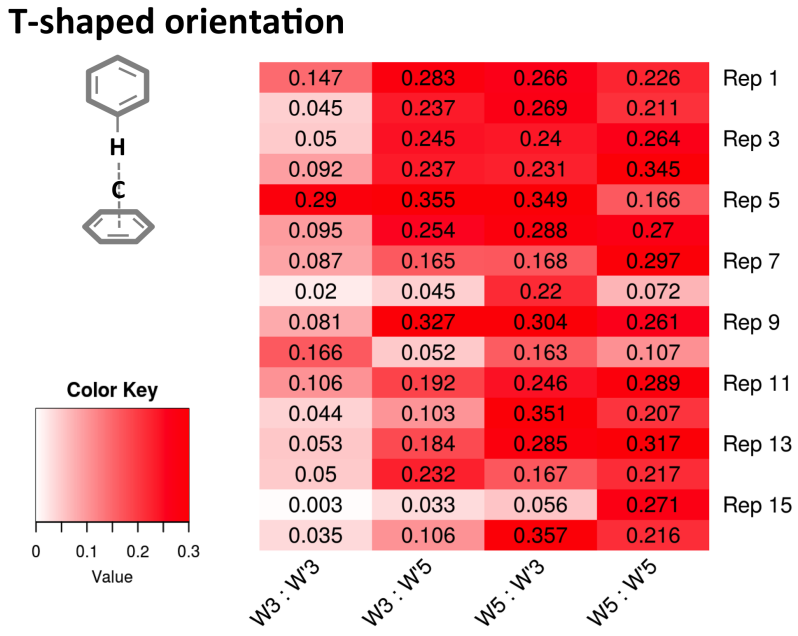


Figure 3.2: Presence of backbone hydrogen bonds between two facing antiparallel RRWTWE chains. The pairs highlighted in blue, green and pink corresponds to three different chains arrangements, shown in the schemes below the histogram. Bonds are labelled by amino acid pairs, named as in the schemes and underlined with matching color; occupancy is averaged over 16 simulations of 20 ns each.



(a)



(b)

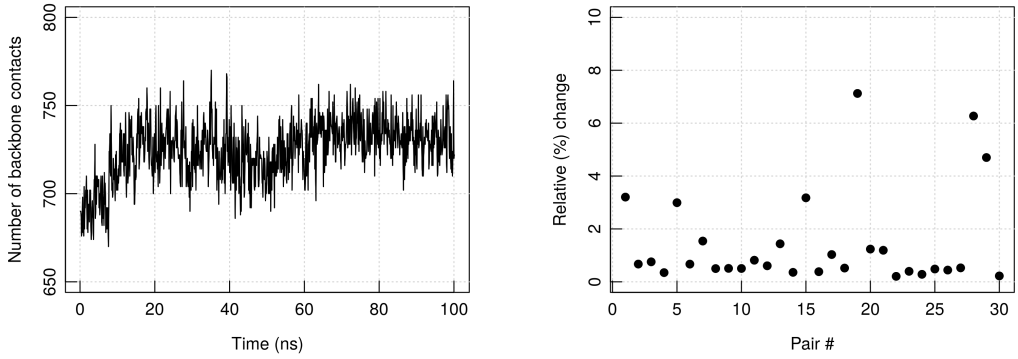
Figure 3.3: Presence of parallel or T-shaped π-stacking between each possible pair of facing Tryptophan residues in two RRWTWE chains arranged in an antiparallel β-sheet. For each replica the map gives the fraction of time for which a parallel or T-shaped π-stacking interaction has been observed (total simulation time 320 ns).

rangement is favoured.

The hydrophobic interactions between Tryptophan residues (in white in Figure 3.1, B) results in the creation of a hydrophobic patch which includes four of them on one side of the β-sheet plane. This generates an amphiphilic structure where the hydrophobic core is segregated from the remaining charged residues distributed around. The combination of two stacking β-sheets, paired to match their hydrophobic patches, constitutes an effective supramolecular assemblies to reduce the solvent exposure of such residues (Figure 3.1, D).

Oligomer simulations This pairing strategy, however, needs to be applied in the context of assembly of complete molecules, i.e. three arms joined by the central scaffold. From now on, we will denote the central scaffold as RKGB residue. This was the name given to the moiety by the ATB server [Malde et al., 2011; Koziara et al., 2014] used to compute its atomistic parameters (see Section 3.2). The quasi three-fold symmetry of capzip suggests a regular geometric arrangement; at the same time simulations of a single molecule in solution highlight its flexibility, proposing that multiple arrangements can be accommodated by the molecule. The best examples of geometrically organised protein structures can be found in viral capsids, which are composed by the regular repetition of highly symmetric protein subunits. Inspired by this, we tested whether a geometrical organisation can represent a stable capsule, choosing as minimal representative geometry a truncated icosahedron (buckyball, Figure 3.1, E). This shape has 12 pentagonal faces and 20 hexagonal ones.

Preliminary atomistic simulations (100 ns) were run on a pentagonal subunit (Figure 3.1, C), where ten antimicrobial molecules are arranged in two stacking pentagons. Capzip arms are paired in antiparallel β-sheets to form each pentagon, and the two of them are facing with their Tryptophan residues in contact (Figure 3.1, D). The simulations proved the cohesion between molecules belonging to the subunit. Specifically, the number of contacts between backbone C^{α}s augmented slightly in the first 20 ns (Figure 3.4(a)), due to the compaction of the unpaired external arms toward the core of the structure (see Supplementary Figure 3.40), which made the Coulomb energy decrease. Moreover, for each pair of facing chains (see definition in Figure 3.4(b)), we computed



(a) (b)

Figure 3.4: (a) Number of backbone contacts during a simulation of a pentagonal subunit. (b) Variability of the inter chain average distance between the 30 facing chains in the pentagon. The 30 pairs defined as facing are the chains belonging to the same β-sheet (5 for each of the two stacking pentagons), and for each stacking β-sheet the 4 possible inter-pentagon (inter-layer) pairs of chains.

the distance between their centres of mass. Figure 3.4(b) reports the variance of these distances normalised by their average value, as a measure of the cohesion of the subunit, showing that in the majority of the cases less than 2% of variability is observed. Overall, the block is not rigid, and does indeed deform and go back to the pentagonal shape multiple times in the simulation, but it is keeping the original pairing of the molecules.

Capsule simulations The pentagonal subunit respects the building principles of β-sheet pairing between antimicrobial sequences. Moreover, the double layer structure allows to partially screen the hydrophobic residues from the solvent, as they are in favourable contact with each other (Figure 3.1, D). Twelve of these units can be used to build a truncated icosahedron (Figure 3.1, E): they constitutes its pentagonal faces, while the hexagonal ones are formed when the arms extending from the pentagons join together.

Each capzip molecule is centred in one vertex of the polygon, with the arms laying alongside the edges departing from it. On each edge two arms coming from opposite sides meet in an antiparallel fashion. The full truncated icosahedron (Figure 3.1, E) has two concentric layers, for a total of 120 molecules, and initial radius of 7.7 nm. Thus it will be called *buckyball bilayer* in the following. As anticipated, this geometry

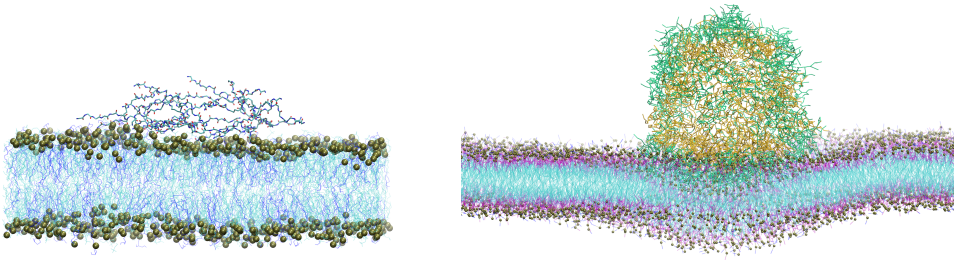
represents a minimal model of the possible structures capzip adopts in solution.

The structure was simulated at atomistic and coarse-grained levels, respectively with the GROMOS 53A6 [Oostenbrink et al., 2004], SIRAH [Machado et al., 2019] and MARTINI [Marrink et al., 2007; Monticelli et al., 2008] force fields (with both standard and polar water [Yesylevskyy et al., 2010]). From the final configurations of the MARTINI coarse-grained model (standard water), atomistic coordinates were obtained and simulated, to be compared with the original atomistic dynamics. Moreover, additional simulations were run at all the coarse-grained levels on a buckyball monolayer (i.e. built from pentagonal subunits made by one pentagon only), to prove whether the bilayer is more energetically favoured.

Finally, a test study on self-assembly was performed with one of the coarse-grained representations (MARTINI with standard water), starting from capzip molecules randomly placed in solution.

Capzip-membrane simulations A multiscale analysis is needed also to investigate the antimicrobial activity. Being highly costly to simulate the buckyball bilayer on a membrane at the atomistic level, the pentagonal subunit employed to build the complete structure (Figure 3.1, C) was taken as representative of the latter. It was simulated parallel to the membrane plane, with its center of mass at 1.5 nm from the phosphate plane (Figure 3.5(a)), to avoid spending time in sampling conformations with the peptide far from the membrane. This, together with a tailored use of an applied electric field (see Section 3.2), speeds up simulations considerably. To observe the natural binding of the peptide to the membrane, the process of the buckyball bilayer approaching a model membrane was simulated with a MARTINI coarse-grained description (Figure 3.5(b)). The use of the MARTINI polar water [Yesylevskyy et al., 2010] allowed for the introduction of an external electric field, to compare the coarse-grained simulations with the atomistic ones which employ analogous conditions.

Two membrane compositions were simulated for both resolutions, a model bacterial and a model mammalian membrane, to identify the different interactions with the peptide. The first one presents 25% of anionic lipids (DLPG), and the remaining zwitterionic (DLPC), while the second has only DLPC lipids. The choice of the bacterial



(a) (b)

Figure 3.5: (a) Atomistic structure of a pentagonal subunit on a 740-lipid bacterial model membrane of composition DLPC/DLPG 3:1 (initial configuration). Peptide backbone in line and cartoon representation; lipids in cyan lines (DLPC) and blue ones (DLPC), all lipids phosphate in golden van der Waals beads. (b) coarse-grained (MARTINI) representation of the buckyball bilayer on a 2880 lipids bacterial model membrane (final configuration of the trajectory). Protein in bonds representation: green outer buckyball layer, yellow inner one. Lipids in line representation, coloured by bead type, and lipids phosphate in golden van der Waals beads. [VMD software Humphrey et al. [1996]]

model was dictated by the experiments performed on capzip [Castelletto et al., 2016], and the mammalian one was built with the same zwitterionic lipid as for the bacterial to simplify the comparison, e.g. to have membranes with comparable thickness and mechanical properties.

3.2 Simulations details

All simulations were performed with the GROMACS software, version 5.5 and 2016 [Berendsen et al., 1995; Abraham et al., 2015, 2018].

3.2.1 Atomistic simulations

Simulations in solution The atomistic coordinates for the peptidic supramolecular assemblies described in Section 3.1 (Figure 3.1, E) were built combining GROMACS tools and the MOE software [Chemical Computing Group ULC, 2018]. Simulations were run with the GROMOS 53A6 force field [Oostenbrink et al., 2004]. Parameters for the central residue connected with the peptidic chains were computed with the ATB server [Malde et al., 2011; Koziara et al., 2014]; the ones for the bonds joining this residue to the antimicrobial side chains were derived from the values tabulated in

the force field for analogous ones. A python module has been designed to manipulate GROMACS topology objects for the GROMOS force field and add peptide bonds at the location required (see Appendix A.3 for details). This is useful for multibranched peptides which are not supported by the standard GROMACS tools, and for which a manual implementation would be otherwise needed.

The systems were solvated with single point charge (SPC) water [Berendsen et al., 1981] and counter ions were added (Na^+ or Cl^-); further ions were introduced to reach the concentration of 150 mM, to reproduce the experimental conditions. For simulations of a single β-sheets, or of the pentagonal subunit in solution, the systems were energy minimised with a steepest descent algorithm, then equilibrated in the NVT ensemble with decreasing positional restraints at increasing temperatures (100 K, 200 K, 250 K, 300 K and respectively 1000, 1000, 500, 250 kJ/mol·nm^2 restraints, 100 ps each); then in the NPT ensemble, without restraints, at the same temperatures steps and for the same time. Production followed for 100 ns.

For the buckyball bilayer structure the above equilibration was insufficient. Due to the construction procedure, two thirds of the arms are not properly paired along the edges, and with a short equilibration they drift away from each other within a few nanoseconds. Therefore, after an NVT equilibration as above, strong flat-bottom restrains (1000 kJ/mol·nm^2) were placed between the center of mass of imperfectly aligned arms throughout the NPT heating, to penalise their mutual separation with respect to their initial distance (100 ps runs at 100 K, 200 K and 250 K and 35 ns at 300 K). This was followed by a series of 10 ns runs at 300 K with decreasing flat-bottom restraints (750, 500 and 250 kJ/mol·nm^2) and by a free production run (100 ns). Three different replicas were run, generated from the final configuration of the 300 K NPT run with 1000 kJ/mol·nm^2 restraints.

Throughout all the simulations, the temperature was maintained by independently coupling the protein and the solvent (plus ions) to two external temperature baths using a velocity rescale thermostat [Bussi et al., 2007] with coupling constant τ_T of 0.1 ps. The pressure was kept at 1 bar by Berendsen [Berendsen et al., 1984] or Parrinello-Rahman barostat [Parrinello and Rahman, 1981] (for the equilibration phases and the production run respectively) using an isotropic coupling, with isothermal compressibility of 4.5 $\times$

10^{-5} bar^{-1} and coupling constant τ_P of 1 ps. Electrostatic interactions were treated using the smooth Particle Mesh Ewald (PME) algorithm [Essmann et al., 1995] with a Fourier grid spacing of 0.12 nm, and a short-range cutoff of 0.9 nm. The van der Waals interactions were treated with a plain 0.9 nm cutoff. All atomistic runs were performed using a 2 fs time step. An overview of the simulations of peptide assemblies in solution is given in Table 3.1.

Simulation with membranes The atomistic coordinates for the bacterial membrane bilayer were built with the PACKMOL software [Martínez et al., 2009], from pdb files of a single DLPC [Poger et al., 2010] and DLPG [Kukol, 2009] molecule. Two bilayers were built, made respectively of 512 and 740 lipids, with composition DLPC/DLPG (3:1). The initial area per lipid was set to 0.70 nm^2, above the values found experimentally for either lipid species (0.608(12) nm^2 for DLPC [Kučerka et al., 2011] and 0.656(12) nm^2 for DLPG [Pan et al., 2012]). An area per lipid compatible (equal within the error) with the experimental values was reached during a 400 ns equilibration (see details below), the final configuration of which was used for simulations with the peptide. Given the analysis performed on the bacterial bilayer (see Section 3.5.1), for DLPC we opted for simulating a large membrane only. Accordingly, we produced a bilayer with 748 lipids and initial area per lipid of 0.68 nm^2.

For membrane-peptide atomistic simulations, the initial configuration was generated from the equilibrated bilayer and the equilibrated pentagonal subunit (after 100 ns run with positional restraints on the C^α), placing the subunit plane parallel to the membrane and close to it (Figure 3.5(a)). The inflategro script [Kandt et al., 2007] was used to solve the partial overlap of peptide side chains with lipid molecules, removing the overlapping lipids if necessary. The two bilayers fit the pentagonal subunit with respectively 3.5 nm and 5.4 nm distance between its periodic boundary images (along both x and y).

For simulations involving membranes, the version 54A7 of the GROMOS force field [Schmid et al., 2011] was initially chosen and it is thus used for simulations of the 512-lipid bilayers, but, upon further research, version 54A8 [Oostenbrink et al., 2005; Reif et al., 2013] was deemed more suitable and thus selected for the runs on the larger

membranes. Lipid parameters were taken from [Poger et al., 2010] for DLPC, while for DLPG they were built from the ones available in the literature for POPG [Kukol, 2009].

The simulations set-up is as above, except for the use of three thermal coupling groups (peptide, membrane, water plus ions), a semi-isotropic pressure coupling, and a larger cut off radius for both Coulomb and van der Waals interactions (1.2 nm). Additionally, for the 512-lipid membrane, a Reaction Field [Tironi et al., 1995] was used instead of PME long range electrostatic treatment (with cut off radius 1.4 nm), inherited from the setup of the simulations used for lipids parametrisation. This was changed to PME when simulating the larger membrane, to be more consistent with the protein parametrisation. Control simulations on membrane bilayers without peptide showed that the results in terms of area per lipid are compatible with the ones obtained using PME.

Each membrane bilayer was first equilibrated for 50 ps in NPT conditions at 50 K, then the temperature was gradually increased up to 300 K in 500 ps, and finally a 400 ns production was run. A similar equilibration procedure was followed for peptide-membrane systems.

Additional simulations were performed applying an external electric field to the membrane, pointing from the side hosting the peptide to the opposite one, to mimic the membrane potential and verify how the peptide affects the response to external stimuli. In a first run the field was increased by 20 mV/nm steps every 200 ns (or 10 mV/ns when reaching the critical value), until poration was induced (at 130 mV/nm). Another simulation was performed for both the 512 and 740-lipid bilayers with the peptide, with the threshold field of 130 mV/nm, starting from the unperturbed membrane configuration (three replicas each). As a control, analogous test simulations were run on the 512-lipid bacterial bilayer without the peptide, assessing the electroporation threshold at the higher value of 140 mV/nm.

To be noticed that the field across the bacterial inner membrane can be estimated to be around 35 mV/nm, and across the mammalian one around 20 mV/nm (an estimate computed from, respectively, a -130/-150 mV and -70/-90 mV potential [Yeaman and Yount, 2003; Wilson et al., 2011] and an estimate membrane thickness of 4 nm).

However, previous computational work often explored the effects of higher fields, up to 500 mV/nm [Tieleman, 2004; Böckmann et al., 2008; Piggot et al., 2011], to witness poration within the simulations time, according to the resources available.

An overview of the simulations of peptide-membrane systems is given as well in Figure 3.1 (control simulations on pure membrane are listed in SI Table 3.8).

3.2.2 SIRAH coarse-grained simulations

SIRAH coarse-grained simulations were run with the SIRAH force field, version 2.0 [Machado et al., 2019]. Peptide coordinates for the buckyball geometry were obtained from the atomistic ones using the converter distributed with force field, with a customised mapping for the central residue. Parameters for the central residue were built from comparison with similar chemical moieties. All simulations were run adding Cl^- counter ions to balance the positive charges of the peptide and additional Na^+ and Cl^- ones to reach a 150 mM concentration.

While for simulations of the peptide buckyball in solution our multiscale procedure includes SIRAH, for simulations on membranes we focussed on atomistic and MARTINI simulations only. This has been performed in the interest of time, and due to technical difficulties in preliminary SIRAH runs with membranes (in particular in tuning the pressure coupling). This is likely due to a suboptimal equilibration procedure; indeed, there are few benchmarks on SIRAH for lipids so far [Barrera et al., 2019], and none for systems as large as the one studied here.

For SIRAH simulations, the temperature coupling was performed with a velocity rescale thermostat [Bussi et al., 2007] and coupling constant τ_T of 0.1 ps, and the pressure coupling at 1 bar pressure, with 4.5×10^{-5} bar^{-1} isothermal compressibility, using a Parrinello-Rahman barostat [Parrinello and Rahman, 1981] with a τ_P of 6 ps. Electrostatic interactions were treated using the PME algorithm [Essmann et al., 1995], with a short-range cutoff of 1.2 nm and relative dielectric constant of 1. The van der Waals interactions are treated with a 1.2 nm cutoff and no long range corrections.

After energy minimization, a 4 ns NVT equilibration was run at 300 K, followed by a 10 ns NPT run, both with positional restraints (1000 kJ/mol·nm^2) on the solute. Two 10 ns run (NPT ensemble, 300 K) were then performed with backbone restraints of

1000 and 100 kJ/mol·nm^2, respectively. Similar to the procedure adopted for atomistic simulations, flat bottom positional restraints (1000 kJ/mol·nm^2) were enforced on the unpaired arms during the latter. After this, three additional 10 ns equilibrations were run at 300 K, with no backbone restraints and decreasing flat bottom ones (respectively 750, 500 and 250 kJ/mol·nm^2). Finally the production run was carried on for 1 μs. All runs were performed with a 20 fs time step.

3.2.3 MARTINI coarse-grained simulations

For MARTINI [Marrink et al., 2007; Monticelli et al., 2008] coarse-grained simulations, peptide coordinates were obtained from the atomistic ones using martinize.py [de Jong et al., 2013] with a customised mapping for the central residue. Parameters were obtained with the same script for the arms, while pycgtool.py [Graham et al., 2017] was used for the central residue. Parameters for the joining bonds were derived from tabulated values of analogous ones.

Initial coordinates for the self-assembly simulations were obtained with GROMACS tools, placing the desired amount of coarse-grained molecules in the simulation box with random positions. The side of the cubic box was chosen as 44 nm, and three different peptide concentrations were simulated: 1.25 mM (60 molecules), 2.50 mM (120 molecules) and 10 mM (480 molecules), all for 10 μs.

The bacterial and mammalian model membranes, hosting 2880 lipids each, were built with insane.py [Wassenaar et al., 2015], with composition DLPC/DLPG (3:1) and pure DLPC respectively. The simulations parameters used for lipids are consistent with Marrink et al. [2004]. The peptide-membrane systems were built placing the buckyball at a minimum distance of 1 nm from the membrane surface.

For all standard MARTINI simulations, counter ions only were added, while for the ones run with Polar MARTINI, additional Na^+ and Cl^- were inserted to reach a 150 mM concentration.

For simulations with the standard water model, the temperature coupling was performed with a velocity rescale thermostat [Bussi et al., 2007] with a coupling constant τ_T of 1 ps. An isotropic or semi-isotropic pressure coupling was applied (simulations of peptide in solution and peptide on membrane respectively) at 1 bar pressure, with 4.6

$\times 10^{-5}$ bar^{-1} isothermal compressibility, using a Berendsen [Berendsen et al., 1984] or Parrinello-Rahman barostat [Parrinello and Rahman, 1981] (equilibration and production phase respectively) with τ_P of 2 ps or 12 ps. Coulomb interactions were treated with a Reaction Field scheme [Tironi et al., 1995] and cut off radius of 1.1 nm, van der Waals interactions with a cut off scheme and the same cut off radius. The relative dielectric constant is set to 15. Simulations performed with the polar water model were run with the parameters above, except the relative dielectric constant set to 2.5, and the choice of a PME scheme for the long range Coulomb interaction (1.2 nm cut off radius). The choice of the respective dielectric constant and electrostatic treatment is in agreement with the optimal setup found in the parametrisation publications of the respective water models.

For simulations of the capsule in solution, for both water models, after energy minimization, four 10 ns equilibration runs (NPT ensemble, 300 K, 10 fs time step) were performed with flat bottom positional restraints on the unpaired arms (respectively at 1000, 750, 500 and 250 kJ/mol·nm^2) - as done for the other force fields. Finally the production run was carried on for 1 μs. All runs were performed with a 20 fs time step.

It is interesting to notice that the MARTINI force field with the standard water model produced very similar results with or without such refined equilibration procedure, as confirmed by simulations of the buckyball run after a short unrestrained equilibration only (500 ps, 1 fs timestep). However, we choose to follow the same procedure as for the other force fields to have more consistent results, and be certain that the differences observed are due only to the parametrisation an not to the equilibration procedure.

From the final configurations of two out of the three standard MARTINI replicas, atomistic coordinates were obtained using the MARTINI backward tool (version 5) [Wassenaar et al., 2014] and run for additional 200 ns with the set up used for atomistic simulations.

The equilibration for the self-assembly simulations (run only with the standard MARTINI model) included only an energy minimization, followed by a 1 ns equilibration with the Berendsen barostat, before switching to Parrinello-Rahman for the 10 μs production.

The membranes used in MARTINI simulations are equilibrated for 1 μs with standard water and the final configuration is used to build the peptide-membrane systems, together with the capsule structure (obtained after the equilibration runs). The full system is energy minimised, equilibrated for 500 ps and production is followed for 10 μs for simulations with the standard water, and 500 ns with polar water (due to the faster binding and the speed up granted by the electric field).

The adoption of polar water allows to perform electroporation experiments also with the MARTINI force field (while the standard water, not bearing a dipole, is unable to screen the externally applied electric field, resulting in unphysical effects). We thus resorted to a procedure similar to the one employed for atomistic simulations, testing an external electric field of magnitude 20 mV/nm and 40 mV/nm, without proceeding further as the latter gave poration in presence of the peptide. It is possible that longer simulations would allow to observe this behaviour even with lower values of the force field, however, in the interest of time, we selected these two for further investigation.

We selected a configuration from the early stages of poration (at 40 mV/nm), where the pore showed a diameter of 2 nm, and prolonged the simulation switching to isotropic pressure coupling. This allows the pore to expand in a controlled manner. Indeed, in a semi-isotropic pressure coupling scenario, the pressure in the z direction suddenly decreases when a pore forms, as water can pass through the membrane. This is also the direction of the applied electric field. The barostat algorithm reacts decreasing the z side of the box, in an effort to restore the target pressure in this direction, but to accommodate all the molecules the box must necessarily expand in the x and y directions in an excessive way.

Control simulations with the two values of the electric field mentioned were performed on the membrane alone. The same procedure, i.e. simulations with capsule and electric field plus control simulations on the membrane only was applied to the pure DLPC membrane as well (except for the additional simulation after poration, as this is not observed for the DLPC membrane).

Table of simulations of capzip assemblies

Peptides	Lipids	Box (nm)	C (mM)	E (mV/nm)	Time (ns)	Rep.
United atom GROMOS (GR)						
β-sheet	–	4	0	–	20	16^a
PS (10)	–	12	0	–	100	1^a
BB (120)	–	22	150	–	100	3^a
BB-back (120)	–	22	150	–	200	2^a
PS (10)	512 (b)	12	150	–	500	2^b
PS (10)	740 (b)	14	150	–	500	1^c
PS (10)	740 (m)	14	150	–	500	1^c
PS (10)	512 (b)	12	150	130	75^P, 20^P, 71^P	3^b
PS (10)	740 (b)	14	150	130	60^P, 50^P, 70^P	3^c
PS (10)	740 (m)	14	150	130	20^P, 28^P, 39^P	3^c
Coarse-grained SIRAH (SI)						
BB (120)	–	22	150	–	1000	3
BM (60)	–	22	150	–	1000	2
Coarse-grained Polar MARTINI (MA_P)						
BB (120)	–	22	150	–	1000	2
BM (60)	–	22	150	–	1000	2
BB (120)	2880 (b)	30	150	–	10000	1
BB (120)	2880 (b)	30	150	20	500	1
BB (120)	2880 (b)	30	150	40	168^P	1
BB (120)	2880 (m)	30	150	–	10000	1
BB (120)	2888 (m)	30	150	20	500	1
BB (120)	2888 (m)	30	150	40	500	1
Coarse-grained MARTINI (MA)						
BB (120)	–	22	150	–	1000	3
BM (60)	–	22	150	–	1000	1
Random (60)	–	22	0	–	10000	1
Random (120)	–	22	0	–	10000	1
Random (480)	–	22	0	–	10000	1
BB (120)	2880 (b)	30	0	–	10000	2
BB (120)	2880 (m)	30	0	–	10000	1

Table 3.1: Table of simulations of capzip assembly in water and on membranes. Peptides: PS - pentagonal subunit; BB - buckyball bilayer; BB-back - buckyball bilayer backmapped; BM - buckyball monolayer; Random - random initial configuration. In parenthesis, the number of capzip molecules. Lipids: number of lipids and model, with (b) - bacterial model (DLPC/DLPG 3:1), (m) - mammalian model (DLPC). Rep: number of replicas. Force field indicated in bold headers: GR - united atom GROMOS (^{a} version 53A6 [Oostenbrink et al., 2004], ^{b} 54a7 [Schmid et al., 2011], ^{c} 54a8 [Reif et al., 2012]), SI - coarse-grained SIRAH [Machado et al., 2019], MA - coarse-grained MARTINI [Marrink et al., 2007; Monticelli et al., 2008], MA_P - coarse-grained MARTINI with polar water [Yesylevskyy et al., 2010]. Superscript P in the time length denotes observation of poration. For electroporation simulations on pure membranes, see Supplementary Table 3.8.

3.3 Analysis

All the analysis was performed combining tools from GROMACS [Berendsen et al., 1995; Abraham et al., 2015, 2018], MDAnalysis [Michaud-Agrawal et al., 2011; Gowers et al., 2016] and packages developed for the purpose (see A.3).

In Section 3.1 we presented an evaluation of the occurrence of π-stacking between Tryptophan residues in the β-sheet simulations. The computation was performed as follow: for each pair of Trp residues, we computed the minimum distance between the centre of mass of their side chains benzene rings (in function of time), as well as the angle formed by their two planes. We consider as candidates for an aromatic interaction rings which were between 0.45 and 0.7 nm of distance, accordingly to the classical definition [Burley and Petsko, 1986] and what found generally in crystal structures [Anjana et al., 2012]. Then, we defined a parallel π-stacking if the angle between the ring was smaller than 45°, and a T-shaped (perpendicular) one if the angle fell between 80° and 100° (as observed in the majority of T-shaped stacking in crystal structures) [Anjana et al., 2012]. Occupancy was found normalising the number of frames in which the interaction appeared by the number of frames considered for the analysis (2000 $\times$ 16, 1 ever 100 ps).

For the β-sheet simulations and whenever in the following we mention hydrogen bond computations, this was performed using the default set up of GROMACS. With such setup, a donor atom (with a polar hydrogen bonded to it) and an acceptor one are classified as interacting through hydrogen bond if the donor-acceptor distance is smaller than 0.35 nm and the hydrogen-donor-acceptor angle is smaller than 30°.

3.3.1 Simulations in solution

Several analysis on the structure and chemico-physical properties of the simulated buckyball bilayer in solution were performed.

- The Radius of gyration (R_g), and Root Mean Square Deviation with respect to the initial configuration (RMSD) were computed (GROMACS).
- To get the average distribution of the mass of the capsule around its center, the Radial Distribution Function (RDF) of protein masses around their center of mass

was computed (GROMACS), considering only the second half of the trajectory. The profile was fitted with a Gaussian function (R software [R Core Team, 2015]): the position of its maximum can be taken as the average radius of the capsule, and its Full Width at Half Maximum (FWHM) as an estimate of the capsule wall thickness.

- The dynamical character of the structure was assessed computing the correlation of motion between molecules (GROMACS). The central atom (or bead) from which the arms depart was taken as reference. For all the pairs (i, j) of such reference positions, the covariance of motion was computed along each direction. The total covariance was obtained as: $\sigma^2(i,j) = \sigma_x^2(i,j) + \sigma_y^2(i,j) + \sigma_z^2(i,j)$. Finally the measure was normalised to obtain the correlation:

$$corr(i,j) = \frac{\sigma^2(i,j)}{\sqrt{\sigma^2(i,i) \cdot \sigma^2(j,j)}}. \tag{3.1}$$

- The pairing of the arms was quantified as follow. Two arms are defined as paired if their center of mass is closer than a cut off distance of 1.2 nm (GROMACS and R postprocessing). This simple measure discards more precise information on the orientation of the chains with respect to each other, and aims at checking whether the network of molecules present in an ideal buckyball structure is maintained. In the ideal buckyball, contacts within the same layer sum up to 90 for each layer (and so 180 in a bilayer). This measure can be easily applied to any description (atomistic or coarse-grained) without disagreement in the interpretation. Given the loose character of the measure, we considered only the pairings surviving more than 90% of the simulation time.

- To characterise the chemical determinants that promote the assembly, we investigated the interactions between amino acids of different types, computing the number of contacts between backbone and side chains of single amino acids (function implemented in MDAnalysis, see Appendix A.3). We filtered for the ones present at least 50% of the simulation time, and remove the ones between residues in the same arm, finally we classified them by amino acid type. Backbone

contacts are computed considering C^{α}s of residues (or the corresponding coarse-grained bead); side chains ones using selected reference atoms/beads (respectively for GROMOS, SIRAH and MARTINI: CZ/BCZ/SC2 for Arg, CZ2/BNE/SC4 for Trp, OG1/BPG/SC1 for Thr and CD/BCD/SC1 for Glu); finally mixed ones if the proximity is between a C^{α} and the side chain reference position. The distance threshold is set to 0.6 nm.

- For atomistic simulations, the inter molecular hydrogen bonds between amino acids have been computed (as explained previously with GROMACS) and grouped by amino acid type, and by region of occurrence (e.g. between two backbones, side chains or connecting a backbone atom and a side chain one).

- The Solvent Accessible Surface Area was computed for each amino acid (GROMACS). Its value was averaged in time and over all the residues at the same position in the capzip sequence. For atomistic runs, it was also normalised over the reference value for the corresponding residue type X, obtained as the theoretical measure of its SASA in a Gly-X-Gly tripeptide [Tien et al., 2013]. The resulting Q_{SASA} takes into account the size of the side chain, giving a measure of exposure which can be compared between different amino acids. This normalisation is however inappropriate for coarse-grained models, due to the differences between the atomistic SASA obtained in experiments and the beads employed in the simulations, so that only the absolute SASA values are given.

- Finally, the contribution to the total energy of the system deriving from Coulomb and Lennard-Jones interactions were compute (GROMACS), breaking them into interactions between protein components and between protein and solvent.

3.3.2 Simulations on a membrane

Many properties can be extracted from a membrane simulation. We report here the ones selected for analysing the simulations including membranes and capzip, while Chapter 4 gives an extensive overview of more of them, reiterating and expanding the description of the ones listed below.

- We monitored the area per lipid (ApL), i.e. the average space available to each lipid on the (local) membrane plane. In the case of approximately flat membranes aligned to the xy plane, this can be assessed from the product of the lateral dimensions of the simulation box divided by the number of lipids in one leaflet. For highly curved membranes, as the ones obtained under the effect of a strong electric field, this measure does not reflect the true spacing between lipids. In that case the true ApL was computed through the algorithm developed by Braun et al. [2011]: this first identifies the undulating reference surface of the membrane $u(\mathbf{r})$ fitting a Fourier series to a set of reference positions (one atom per lipid, presently the Phosphorus atom or the Phosphate bead for coarse-grained simulations), then it computes the true ApL based on $u(\mathbf{r})$ as:

$$ApL = \frac{1}{N} \int_{x_{box}} \int_{y_{box}} \left[\sqrt{1 + (\nabla u(\mathbf{r}))^2} \right] dx\, dy \tag{3.2}$$

  where the gradient $\nabla u(\mathbf{r})$ is zero for a perfectly flat surface.

- For atomistic simulations, the deuterium order parameters S_{CD} of the acyl chains is a good measure of their relative orientation and whether the chains adopt for it a narrow range of values [Van Lehn and Alexander-Katz, 2014; Douliez et al., 1998; Piggot et al., 2017]. The orientation θ with respect to the outward leaflet normal is computed for each carbon-hydrogen bond in a given position i along the chain, for each lipid in the bilayer. Their spread is evaluated according to the ensemble average:

$$S_{CD}(i) = \frac{1}{2} \langle 3 \cos^2 \theta_i - 1 \rangle. \tag{3.3}$$

  As the GROMOS force field employs a united-atom representation, the tetrahedral positions of the hydrogens are constructed based on the neighbouring carbons positions.

- Another measure of the regular packing of lipids in the xy plane (parallel to the membrane surface) is given by the hexagonal order parameter S_6 [Uppulury et al., 2015], usually employed to quantify the transition to a gel phase. Specifically, each lipid tail chain was represented by its position on the xy plane, computed as the

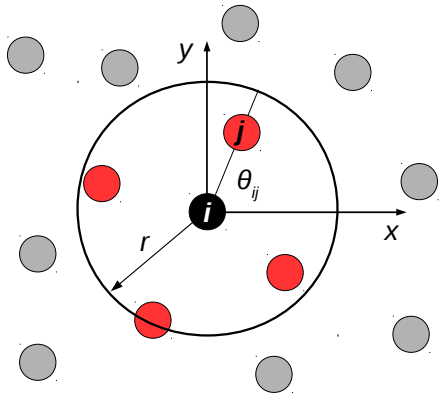


Figure 3.6: Scheme of S_6 computation: in black the reference chain i, in red the chains belonging to the neighbouring chains set $\{n\}$; r was chosen as 0.65 nm.

average x and y position of its carbon atoms. For each chain j the neighbouring chains $\{n\}$ were identified as the ones within a 0.65 nm radius from j (Figure 3.6). Then S_6 is defined as:

$$S_{6,j} = \frac{1}{6}\left|\sum_{k\in\{n\}} e^{6i\theta_{jk}}\right| \tag{3.4}$$

with θ_{jk} the angle between the vector connecting j and k, and the x axis (and i the imaginary unit). A chain is in gel phase if it has an hexagonal order parameter larger than 0.72 [Uppulury et al., 2015].

- To evaluate the mobility of lipids on the membrane plane, for each simulation we extracted the trajectory of the phosphorus atom of every lipid, removing the jumps derived from the periodic boundary conditions treatment, to have a continuous evolution of the positions. The resulting trajectory must be centred to remove the drift of the whole system, and this can be achieved in two ways: either considering the two leaflets separately and centring each of them within the simulation box at each time frame (L_{COM} procedure) or considering the bilayer as a unique object and translating its center of mass at the origin of the simulation box for each time frame (B_{COM} procedure). Figure 3.7 illustrates the procedures in two cases, showing how they differ if a leaflet translates with respect to the other one.

The MSD of each lipid in one leaflet was computed as a function of time: for a lipid i and a given time t, the displacement $r_i^2(t_{start} + t)$ was computed for all t_{start} such that $(t_{start} + t)$ is within the simulated time. The average over the start times and over the lipids (belonging to the same specie and leaflet) gave the MSD(t):

$$MSD(t) = \langle \langle r_i^2(t_{start} + t) \rangle_{start} \rangle_i. \tag{3.5}$$

This computation was performed discarding the first part of the trajectory, the exact amount of which depends on the system, and it was chosen monitoring the area per lipid. The diffusion coefficient D was obtained from a linear fit of the MSD(t), following Einstein equation in two dimensions [Einstein and Fürt, 1956]:

$$\langle r^2 \rangle = 4Dt. \tag{3.6}$$

The fit was performed in the regions which showed a linear dependence. This implied discarding the first interval of the profile, where the behaviour is not linear, and the last one, where the poorer statistics leads to more noisy data. We choose 50 ns for both of them for atomistic simulations and 100 ns for MARTINI ones. This was done for each simulation condition of the respective resolution. As a final step, the D values from the two leaflets are averaged.

As a short hand notation, we will call the D coefficient obtained from a L_{COM} removal as *pure diffusion*, while the one obtained from a B_{COM} removal as *global diffusion*. The former measures the pure random movement of a lipid with respect to its neighbours, the latter is influenced by possible collective movements of one leaflet with respect to the other.

For atomistic simulations of the protein on the membrane, further analysis included the following:

- we analysed the network of hydrogen bonds between the protein and lipids, classifying the interactions by the protein residue type and the lipid species (or lipid moiety) between which they occur. Hydrogen bonds are evaluated as described previously;

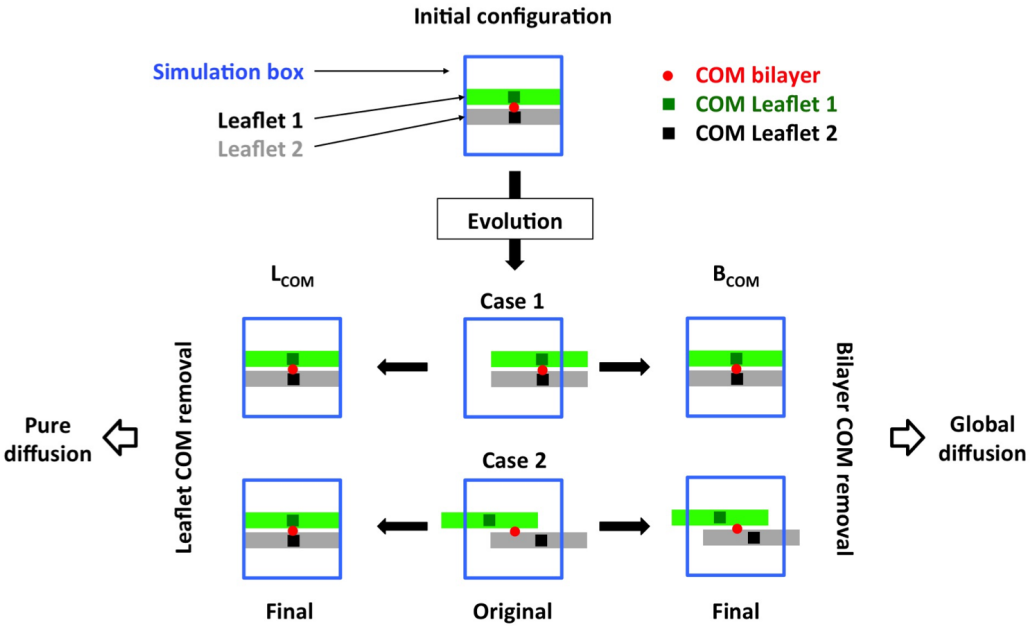


Figure 3.7: Scheme of two different procedures for centring a trajectory of a lipid bilayer: from the *Original* configuration, coordinates are translated to remove either the bilayer or each leaflet movement. The collection of *final* frames, one for each time step, gives the trajectory used to compute the MSD of lipids and thus diffusion.

- for selected simulations the aforementioned diffusion constant was computed selectively on the lipids which, at the initial time, were within a threshold distance from the protein or, conversely, further away from it. In this case, the trajectory was centred with respect to the peptide centre of mass, to understand how lipids move with respect to it;

- the insertion depth of each amino acid in the membrane was calculated as the difference between the z position of the lowest atom of the amino acid and the average of the maximum z coordinate of the five lipids closest to it. This allows to take into account undulations of the membrane by considering the local z coordinate. As for an optimally packed membrane in gel phase each lipids has 6 neighbours, the value of 5 reflects approximately the number of first neighbours in a bilayer in fluid phase.

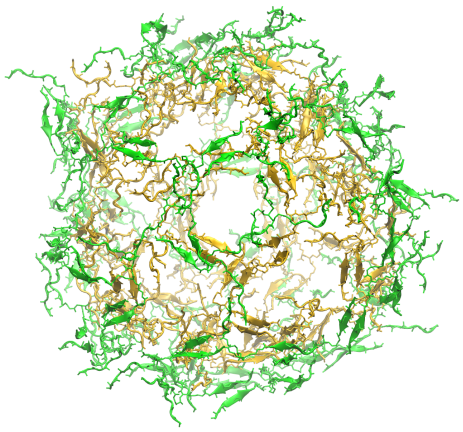


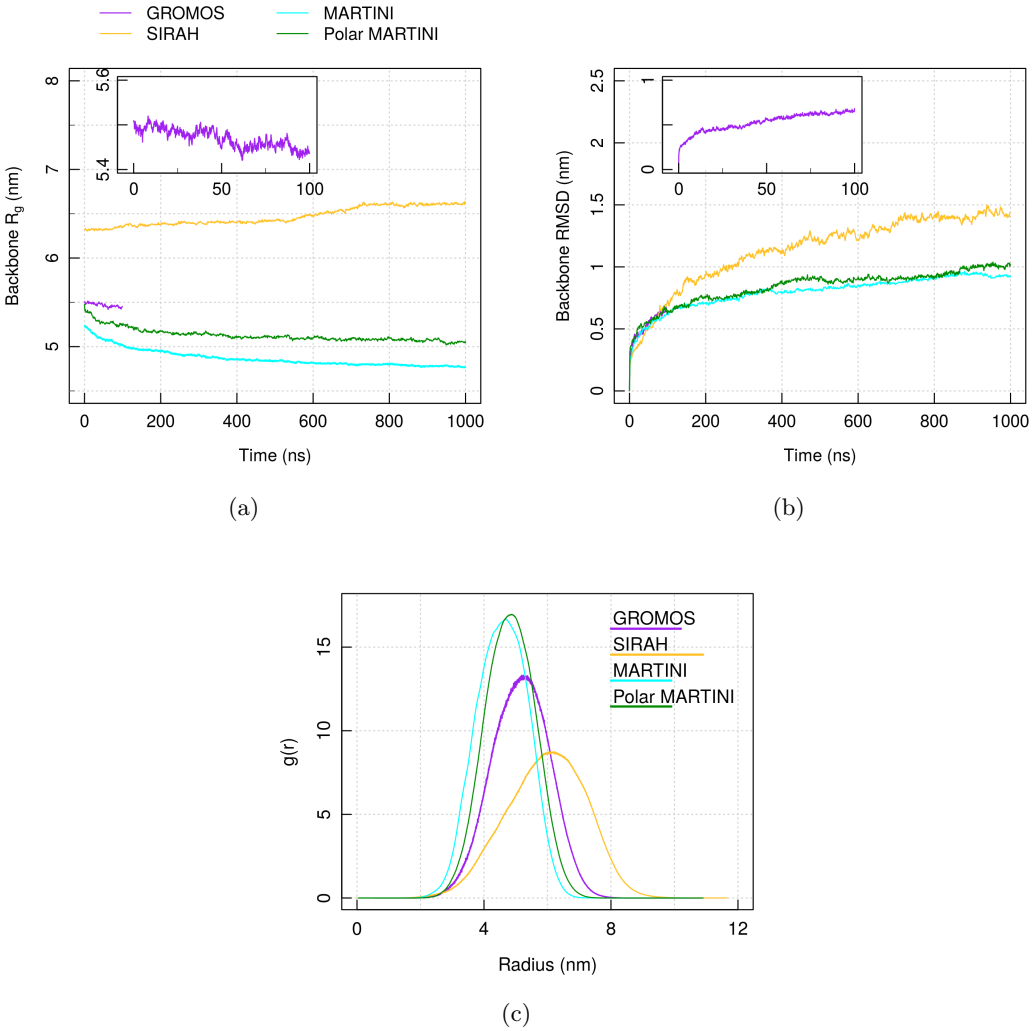
Figure 3.8: Final configuration from an atomistic simulation of a buckyball in solution (100 ns, Replica 1): bonds and cartoon representation, backbone only, green external layer and yellow internal one. [VMD software Humphrey et al. [1996]]

3.4 Results: capsule in solution

We list here the results for simulations of the capsule in solution: starting from the atomistic simulations, we then proceed to compare them with different coarse-grained models. However, when applicable, the plots present the results from all the parametrisations at once, anticipating the discussion of Section 3.4.2, as this makes the comparison easier. In all the plots of the section, atomistic results are color coded in violet, as opposed to yellow (SIRAH), cyan (standard MARTINI) and dark green (Polar MARTINI). Results are shown for Replica 1, unless otherwise specified. The analogous results for Replica 2 are shown in the Supplementary Material (Section 3.7, Figures 3.31-3.39). Supplementary Movies SI_M1, SI_M2 and SI_M3 show GROMOS, SIRAH and Polar MARTINI runs (Replica 1); for the two coarse-grained force fields simulations of the monolayer are displayed as well (see Section 3.4.3).

3.4.1 Atomistic simulations

Global capsule structure Atomistic simulations of the buckyball in solution show a consistently equilibrated structure across the three replicas run (Figure 3.8 and SI movie 1). This is proven by both the stable value of the protein R_g and the almost plateauing backbone RMSD (Figure 3.9(a) and (b)). It is interesting to notice that previous simulations performed with a shorter equilibration, without the phase of employing flat bottom restraints, lead to the immediate disruption of several connections in the



(a) (b)

(c)

Figure 3.9: (a) R_g and (b) RMSD computed on the Protein backbone. Results are displayed for simulations performed in GROMOS (100 ns), SIRAH, MARTINI and MARTINI with polar water (all 1 μs). Inset: zoom on the GROMOS values. (c) RDF of Protein masses around their center of mass, displayed for the same simulation set up as in (a,b). For each label of the legend, the bar has length of the respective FWHM of the Gaussian function fitting the data (thickness estimate). All results are showed for Replica 1 of each simulation set-up.

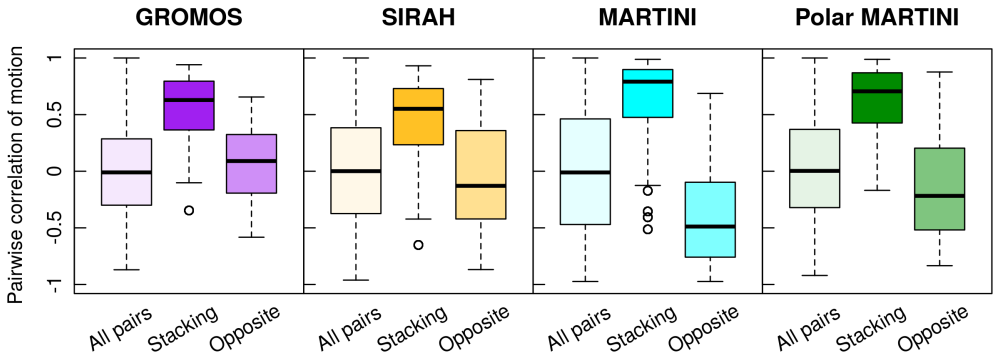


Figure 3.10: Distribution of the correlation of motion between different molecules in the buckyball simulations. Black band: median of the distribution; box: first and third quartiles; whiskers: maximum and minimum, outliers excluded (hollow dots). Results are shown for Replica 1 of each simulation set-up.

buckyball network, resulting in larger R_g. This suggests that the structural pairing present in the buckyball can form only when the chains are in close contact. This is compatible with the long time of assembly observed experimentally (up to 7-10 hours): the closer the chains need to be to trigger the assembly, the rarer this event is statistically happening. Also, this implicitly proves that the self-assembly simulations starting from disordered states will not lead easily to such ordered configurations.

The RDF of the protein masses around the buckyball centre shows no masses nearby the origin (Figure 3.9(c)) and this means that the buckyball remains hollow (given the way RDF is computed and normalised, a uniformly full object would display a flat distribution). A fit of the RDF to a Gaussian curve returns a mean value of 5.1 nm and a FWHM of 2.2 nm, which gives an estimate of the bilayer thickness. A similar computation is repeated for the inner and outer layer separately, providing 1.1 nm of distance between the means of the two distributions. This interlayer distance is compatible with the one between the backbones of stacking β-sheet in structures like densely packed amyloids (1.0 nm [Sunde et al., 1997]).

This thickness value hints at the fact that the two layers are closely packed. This is confirmed by the analysis of the pairwise correlations of motion: molecules at the same polar coordinates (i.e. stacking radially one on the other) have a positive correlation and so move coherently, while the ones at opposite poles, as well as the ensemble of all possible pairs, do not show particular correlation (Figure 3.10).

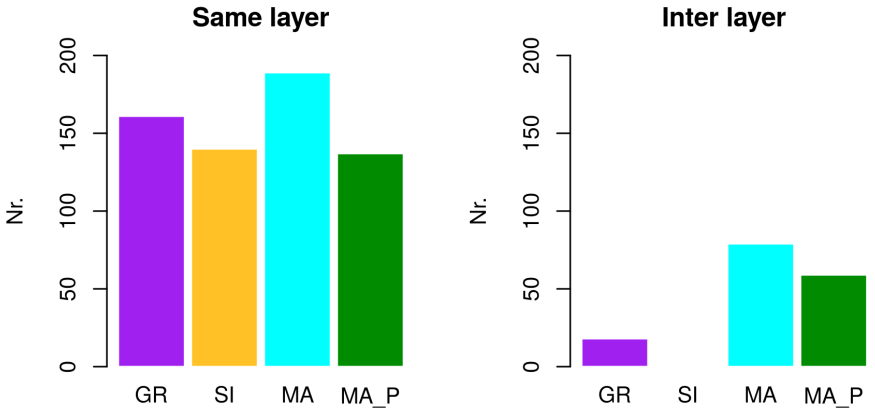


Figure 3.11: Number of paired arms within the same layer and between layers. Cut off distance between arms center of mass equal to 1.2 nm. Only contacts existing more than 90% of the simulation time are counted. Results are shown for Replica 1 of each simulation set-up.

Contacts between arms The measure of how many arms in the buckyball network remain paired for more than 90% of the simulation time is shown in Figure 3.11. An average of 160 pairings is observed between arms belonging to the same layer (summing over inner and outer), and around 20 only for inter-layer ones. This first value correspond to slightly less than 3 pairing per molecule, which would be the value in a perfectly icosahedral structure. This 10% reduction of the arms shows that the structure is not rigid, and some pairings are lost, but nevertheless it maintains the majority of its network of interaction in place. On the contrary, few inter layer contacts are observed within 1.2 nm distance cut off: the arms belonging to two different layers are separated by the space spanned by their side chains, and this keeps the average positions of their backbones at a distance greater than the cut off chosen.

Contacts between residues Figure 3.12 shows the number of contacts between backbones and/or side chains of amino acids which survive more that 50% of the simulation time. They are classified by amino acid type and normalised over the number of residues present for each type (so this takes already into account the fact that there are two Arg and Trp per arm of capzip, but one Thr and Glu only).

The number of backbone contacts per residue is around 1 for Threonine (Thr) and Glutamic acid (Glu), 0.7 for Arginine (Arg), and 1.2 for Tryptophan (Trp) residues (Figure 3.1, B). Thus, on average each residue is paired with another one, except for

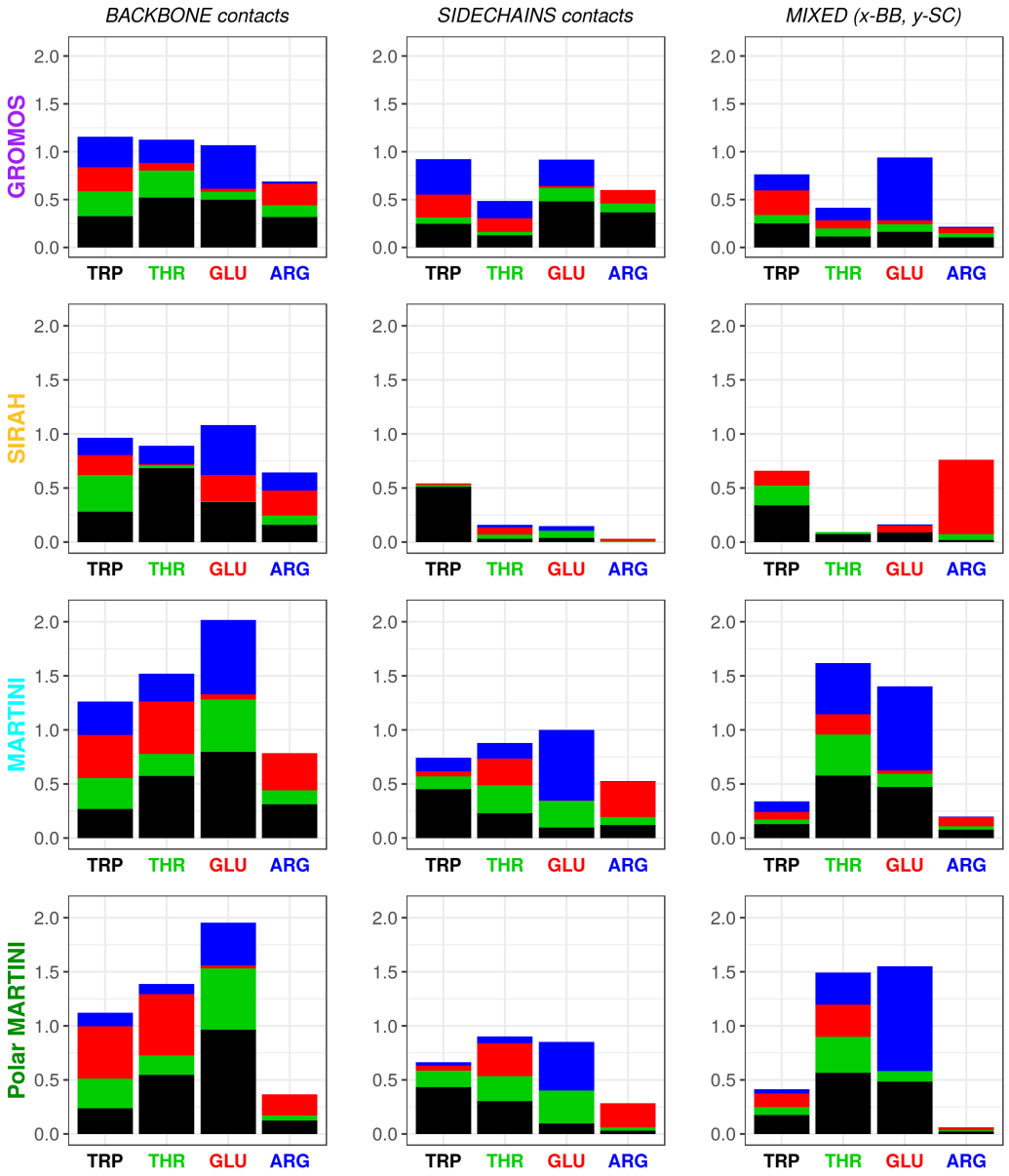


Figure 3.12: Number of contacts per residue type in each arm of capzip: each bar shows the average number for the residue on the x-axis; its color is split by the identity of the partner residue (color coded as in the x-axis legend). For mixed contacts the residue on the x-axis contributes with its backbone. The parametrisation is reported along the y-axis. Results are shown for Replica 1 of each simulation set-up. Only contacts existing more than 50% of the simulation time are considered.

Arginines: only two thirds of them are engaged, likely because some Arg residues are at a terminal positions, and as such they tend to extend in solution, unpaired. On the contrary, Tryptophan is more prone to form contacts, because of its hydrophobic character and central position in the sequence.

The bar plot shows also that there is no strictly fixed arrangement between arms: for example, Tryptophan residues are not paired mainly to Tryptophan ones, as the optimal arrangement would dictate, suggesting flexibility in the structure. Nevertheless, this specific residue has clearly a prominent role in forming contacts with the neighbours both at the backbone level and at the side chains one through cation-π interaction with Arginine (central column of Figure 3.12).

Hydrogen bonds interaction Some of the contacts mentioned above are mediated by hydrogen bond interactions, so the average number of inter molecular hydrogen bonds occurring during the simulations is computed, divided by the number of molecules and classified by residue type. We further separate the hydrogen bonds occurring between backbones and/or side chains. Tryptophan contributes to a large number of backbone hydrogen bonds (Figure 3.13, A), especially with other Tryoptophan residues, consistently with what found in the analysis of contacts carried on previously. Arginine side chains are the most prone to establish H-bonds as a donor with many different amino acid side chains (Figure 3.13, B), but especially with Glutamic acid as expected from the facing positions they occupy in the network arrangement and their opposite charges which attract them closer. Similarly, a common interactions is between Arg side chain (donor) and Glu backbone (Figure 3.13, C), and between Arg backbone (donor) and most of the other side chains (Figure 3.13, D).

Chemical characteristics of the surface: SASA Finally, it is important to understand what residues are exposed at the surface of the structure, especially in view of future applications: in order to make the peptide co-assemble with other products, the two must have a compatible chemical character. The average Solvent Accessible Surface Area (SASA) for every amino acid along capzip arm (Figure 3.14) shows that half of the accessible surface is represented by the charged residues Arginine, while

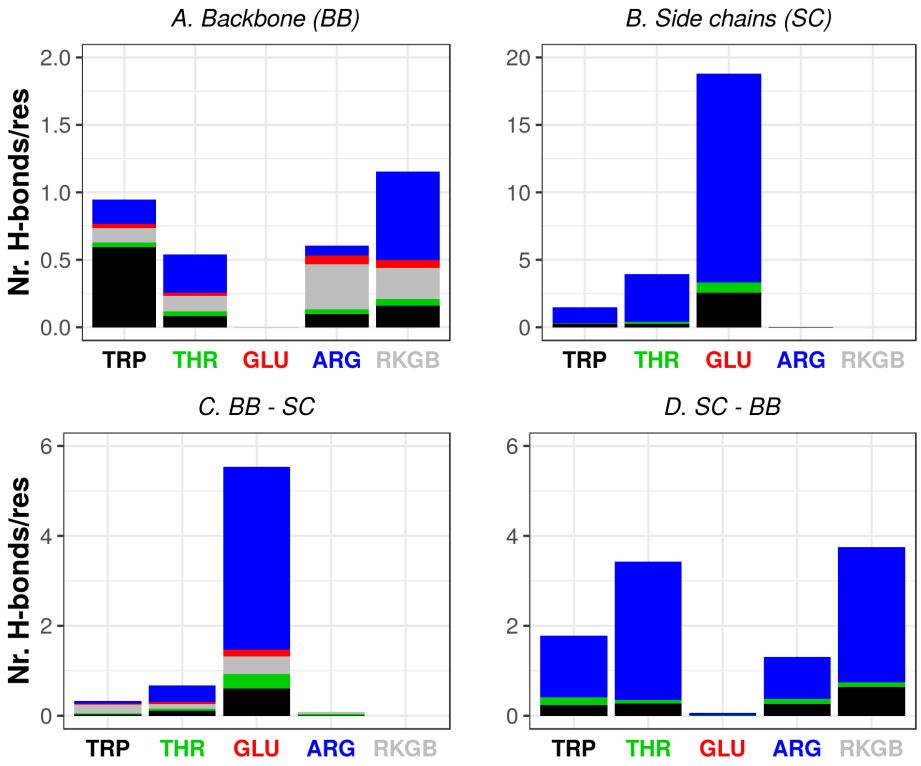
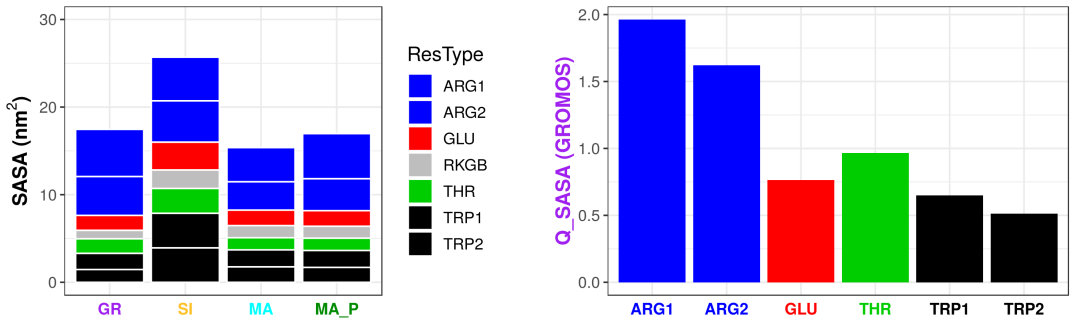


Figure 3.13: Average number of hydrogen bonds per residue occurring between amino acids, including the central scaffold RKGB, for a 100 ns atomistic simulation of the buckyball in solution. Result are shown for Replica 1. For each bar, the residue on the x-axis is the acceptor, and the bar is split by the identity of the donors. In the case of Backbone - Side chain and Side chain - backbone, the first mentioned correspond to the acceptor (and thus the residue on the x-axis).

Tryptophan contribute to it for less than one quarter, despite having bulky side chains. For atomistic simulations, we can compare the SASA of each residue of type X with the reference SASA computed by theoretical modelling of a Gly-X-Gly tripeptide [Tien et al., 2013]. The resulting ratio Q_{SASA} is greater than one for both the Arginines: while for the terminal one (ARG1) it is consistent with the absence of a residue on one of the sides, the fact that also the second has $Q_{SASA} > 1$ proves that these residues are highly exposed in solution. On the contrary, Tryptophan has values around 0.5, due to its propensity to be buried inside the structure, while Glutamic acid and especially Threonine have value closer to one, showing no particular burying.

D-amino acids The results presented above are derived from simulations of the capsule composed by standard amino acid (L-form). As explained in Chapter 1, some experiments focussed on the opposite chiral form (D-version) for stability and immuno-



(a) (b)

Figure 3.14: (a) Solvent Accessible Surface Area (SASA) per molecule, divided by residue types. Results are shown for Replica 1 of each simulation set-up. (b) Normalised SASA over the reference SASA computed for each amino acid type X as the value in a Gly-X-Gly tripeptide.

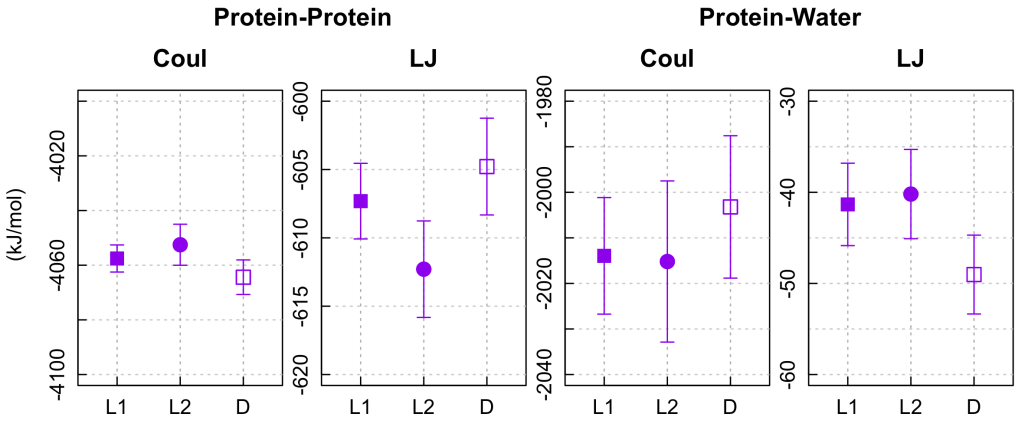


Figure 3.15: Energies of L- and D-amino acids buckyball bilayer. Average values for the second half of the simulation. Two replicas are shown for L-amino acids (L1 and L2), one only was run for D-amino acid (D).

genicity reasons. Experimental work shows similar behaviour for the two enantioners in solution, because they are mirror images of each other.

Accordingly, simulations of D-amino acids should not give any discrepancy with what found before. As a control, we performed one run with such mirror system - and suitable amino acids parameters to keep the D-chirality - finding a similar behaviour of the capsule (i.e. differing from the L- simulations as much as the L- replicas differ among them - Figure 3.15).

This investigation can be performed only at the atomistic level as the coarse-grained models employed represent the C^{α} and the hydrogen attached to it as a unique bead, loosing then the chirality of this centre.

3.4.2 Multiscale comparison of model capsule

We performed a multiscale analysis simulating the capsule structure with different coarse-grained force fields, with a twofold aim: first we wanted to simulate the assembly for a longer time, to observe how its structural properties are maintained on the medium time scale (of the order of the microsecond). Second, we believe that proving the stability of the capsule with different descriptions strengthens the evidence that the assembly proposed is indeed a favourable arrangement of the molecules in solution.

As mentioned before, to this aim we compared simulations run with the SIRAH, MARTINI force fields and MARTINI used in conjunction with polar water (Polar MARTINI). The investigation is also useful to elucidate where the descriptions differ and to infer the advantages of each model. We first comment on the quantities already analysed at the atomistic level (if applicable), and then we extend the analysis to simulations of a monolayer capsule, which has been modelled to verify whether the hypothesis that a bilayer structure is necessary to grant stability was true.

Coarse-grained global structures As foreseeable, the structures obtained with coarse-grained force fields are slightly different among each other and with respect to the atomistic one. The SIRAH bilayer capsule has a structure more expanded with respect to the atomistic one, with a skewed and broader RDF profile, while MARTINI provides a more compact configuration, and finally Polar MARTINI a slightly more expanded structure than standard MARTINI, with a comparable thickness. Respectively, for SIRAH, MARTINI and Polar MARTINI the average radius is 6.0 nm, 4.6 nm and 4.8 nm, with 2.9 nm, 1.9 nm and 1.9 nm average thickness - see Figure 3.9(a) and (c). The difference between the two MARTINI models is likely due to the better properties of solvation of charged groups in the Polar MARTINI description [Yesylevskyy et al., 2010], united with the fact that the buckyball contains a high number of them.

The correlation of motion between molecules for all the coarse-grained force fields is similar to the atomistic one, with a slight anticorrelation between molecules at the opposite poles for MARTINI and Polar MARTINI (Figure 3.10). This is due to the contraction happening at the beginning of the simulation, when the capsule adjusts to the equilibrium size, which depends to some extent on the force field. These effects are

more pronounced for the MARTINI force fields, probably due to the greater cohesion between the beads which makes them moving coherently.

Coarse-grained contacts analysis The number of chains paired in the SIRAH simulations is fewer than in the atomistic ones (Figure 3.11), in line with a more expanded structure, while MARTINI simulations propose a higher number, consistently with the reduced size of the capsule. In particular, the contacts between the two layers are significantly higher in MARTINI. This is expected when the two layers are closer, as suggested by the values of the stacking molecules correlation. Finally, Polar MARTINI agrees with the high number of inter layer contacts of MARTINI, but suggests a partial loss of contacts within the same layer. However, as the structure is compact, the overall shape is still maintained, even without such precise pairing of the arms.

Breaking down the contact analysis by residue type (Figure 3.12), each coarse-grained force field shows a different organisation due to the models of the side chains volumes, which thus occupy a different fraction of the space available. Nevertheless, in all the representations, Tryoptophan has a prominent role in establishing contacts with its neighbours at the backbone level, while different force field disagree on the role of the side chains. Quite surprisingly, the SIRAH force field does not promote interactions between side chains a part from the Tryptophan ones with themselves. This seems due to the more expanded structure of the capsule and consistent with the larger solvation of the amino acids, which are more exposed in solution (see next paragraph discussing SASA values). Finally, none of the coarse-grained force fields seem to capture the preferential cation-π interaction between Arginine and Tryptophan, as can be expected from a less detailed description.

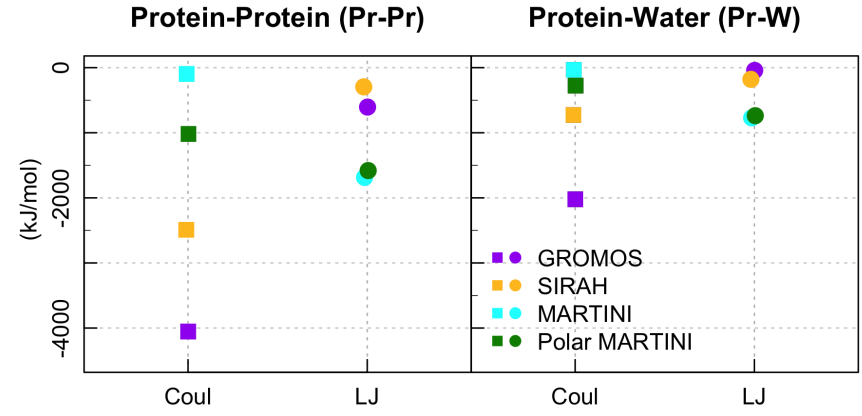
Coarse-grained SASA The values of SASA cannot be compared between force fields, because of the different dimensions of the beads and the number of them employed to describe the residues. However, it is interesting to notice that consistently across force fields, Arginine constitutes around half of the exposed surface, with the exception of SIRAH, where the more expanded structure results in all the residues to be quite exposed (Figure 3.14(a)). This confirms that these charged residues are more

exposed in solution, while Tryptophan ones are buried.

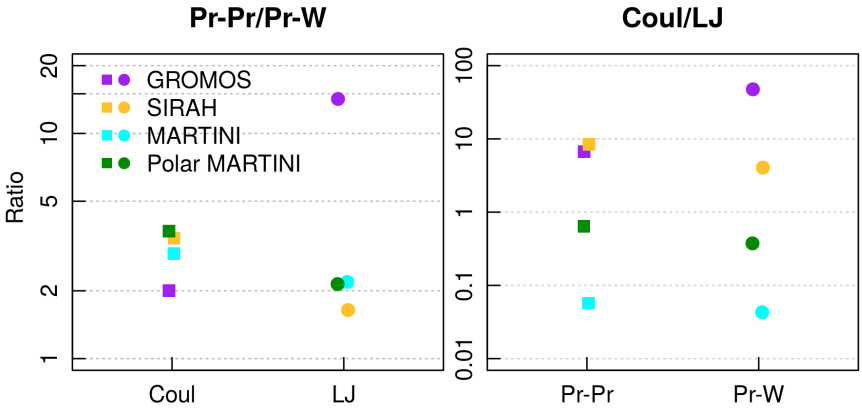
Energetic profile The above results point out that every coarse-grained force field has a particular propensity for an equilibrium distance between peptidic components. This is due to the different solvation property of the water model chosen, and to the balance between the different components of the energy. To better understand this, we computed the Coulomb and Lennard-Jones contribution to the energy due to Protein-Protein interactions or Protein-Water ones. Figure 3.16(a) plots these values averaged over the second half of the simulated time for the atomistic and each coarse-grained force field (for atomistic and SIRAH description the Protein-Protein terms include both the short range interactions and the 1-4 interactions, i.e. the ones computed between atoms separated by three bonds, as they are computed separately during the simulations. For MARTINI models instead they are grouped together in the short range term).

Considering the Coulomb component, it is clear that the mean field approach of standard MARTINI, which consists in adopting a high relative dielectric constant ($\epsilon = 15$) to compensate the absence of water dipoles, decreases sensibly the contribution of the Protein-Protein electrostatics with respect to the two other models. This approach reduces all the Coulomb interactions, while in reality the screening effect due to the water can be seen only, e.g., on distances larger than about 1 nm in a 0.1 M salt solution. This simplified approach instead screens also the contribution of two nearby atoms separated by a distance less than the size of a water molecule. Due to the r^{-6} behaviour of the Coulomb interaction, these short scale contributions are clearly important for the total Coulomb energy. The partial reversion to $\epsilon = 2.5$ performed by Polar MARTINI, together with the introduction of a water dipole, increases the amount of the Coulomb contribution by 10-fold (in absolute value). Interestingly, this is higher than the ratio $\epsilon_{standard}/\epsilon_{Polar} \sim 6$, likely because the reduced short range screening also allows a stronger interactions between opposite charges, bringing them closer and thus contributing more (negatively) to the Coulomb energy.

SIRAH simulations instead are run at $\epsilon = 1$, and present a Coulomb component 2.5 times larger than the one of Polar MARTINI, suggesting that the two models give a similar energy contribution, which is then rescaled by their respective dielectric



(a)



(b)

Figure 3.16: (a) Protein-Protein and Protein-Water non-bonded interactions, normalised per molecule. Values obtained as average on the second half of the trajectory of Replica 1 for each simulations set-up. (b) Ratio between the Protein-Protein and Protein-Water interactions for each force field, for Coulomb and Lennard-Jones respectively; or between Coulomb and Lennard/Jones, for Protein-Protein an Protein-Water interactions separately (note the log scale on y). Values computed as for plot (a). Points are misaligned along x to facilitate the reading.

constant. Finally, the atomistic model ($\epsilon = 1$) suggests even higher (in absolute value) Coulomb energy, roughly the double of SIRAH ones.

These differences in Coulomb energies affects both the Protein-Protein and Protein-Water interactions consistently (Figure 3.16(b), left), but they have consequences on the dynamics because they change the proportion of Coulomb energy with respect to the Lennard-Jones contribution (Figure 3.16(b), right). These interactions are not changed between the two MARTINI models, making them predominant in standard MARTINI, where electrostatics are weak, while they are competing with the Coulomb contribution in Polar MARTINI. On the contrary, SIRAH parametrisation opts for a smaller role of Lennard-Jones with respect to the Coulomb contribution, more consistently with the atomistic description, especially regarding the Protein-Protein interaction. To recapitulate, protein electrostatics have an increasing contribution in MARTINI, Polar MARTINI, SIRAH and atomistic respectively, both in absolute terms and with respect to the Lennard-Jones contribution.

This might partially explain the differences in sizes observed across the models: for example the Protein-Protein Coulomb energy is more negative in SIRAH. However, with a smaller dielectric constant with respect to MARTINI, these contributions are less screened, thus the many positive amino acids composing the capsule (giving a positive net charge) can repel each other more effectively.

Comparing the two MARTINI models instead, Polar MARTINI has a greater Protein-Protein Coulomb component with respect to standard MARTINI (in absolute value, which means a more negative contribution). From this can be deduced that the slightly more expanded structure observed in Polar MARTINI is likely due to a different balance of electostatic versus Lennard-Jones, rather than to the more favourable solvation due to the polar water model. Indeed, the balance between Protein-Protein and Protein-Water components (Figure 3.16(b)) is in the same range for all the three coarse-grained force fields, and it is slightly higher than the atomistic value for Coulomb, while for Lennard-Jones it is lower by about 15-fold.

An interesting follow up on this topic would be investigating whether a tuning of the dielectric constant used in the two MARTINI models can produce more consistent results between them. However, the choice of the constant was optimised to reproduce

at best the properties of bulk water and solvation free energies of ions for both cases. This then raises the question whether the protein parametrisation performs equally good in conjunction with both model, or whether one reproduces better the atomistic Protein-Water interactions.

3.4.3 Bilayer versus monolayer: coarse-grained simulations

To prove also at the capsule level that the bilayer structure is indeed essential to grant a structure which does not disassemble or changes shape, we performed simulations of a monolayer capsule (specifically taking the external layer of the capsule already simulated - Figure 3.1, E) in the three coarse-grained force fields employed so far (Supplementary Movies SI_M2 and SI_M3 show SIRAH and Polar MARTINI runs for the monolayer as well - Replica 1).

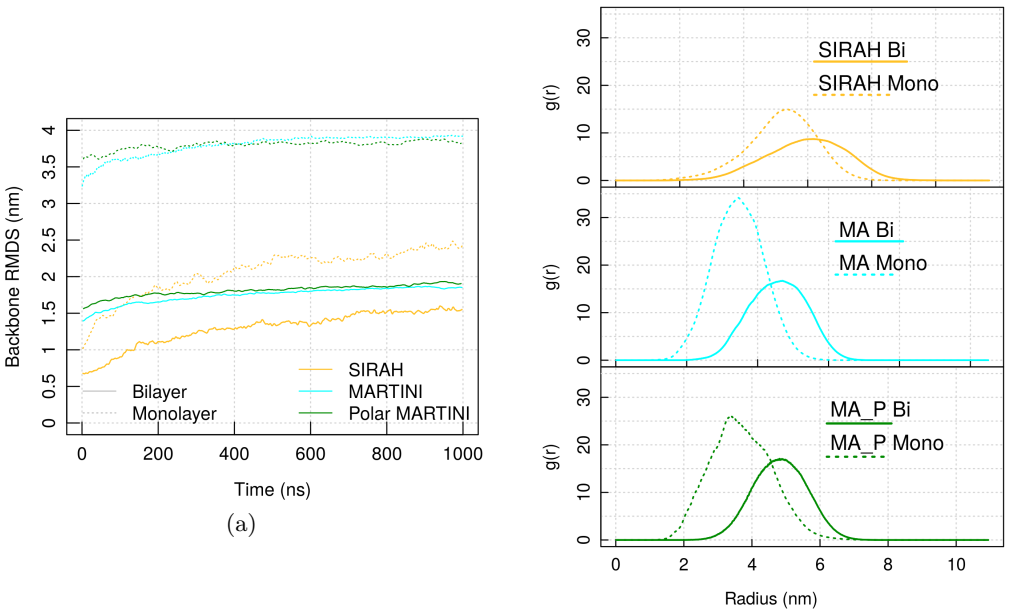


Figure 3.17: (a) RMSD of the monolayer and bilayer structures for SIRAH and Polar MARTINI force fields, with respect to the initial geometrical configuration (external layer of Figure 3.1, E). (b) RDF of Protein masses around their center of mass. For each label of the legend, the bar has length of the respective RDF FWHM (thickness estimate). Results are shown for Replica 1 of each simulation set-up.

The RMSD with respect to the initially built structure (the geometrically regular polyhedra as in Figure 3.1, E) shows that the monolayer undergoes a larger conformational change than the bilayer in the SIRAH force field and the effect is even more

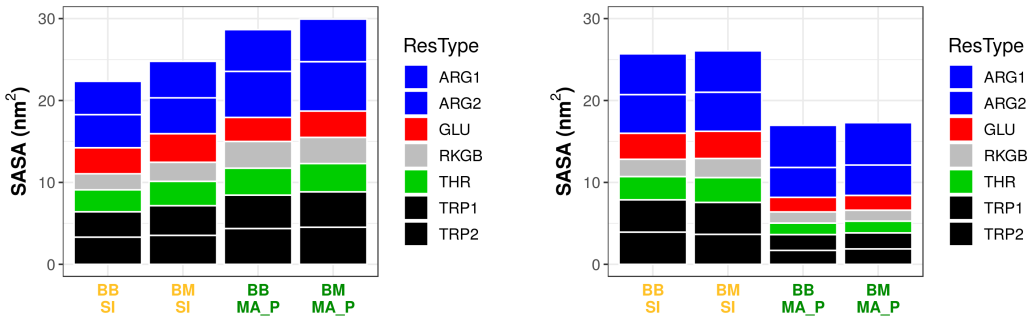
pronounced for the Polar MARTINI (Figure 3.17(a)). As a note, we take as reference structure the regular geometry and not the first frame of the production as major rearrangements happen for the monolayer already in the equilibration phase, and those are different for each force field. This larger change of the monolayer is due to a larger contraction of the structure, which collapses more toward its center (Figure 3.17(b)).

The SASA of each residue type computed on the initial configuration is slightly higher for the monolayer than the bilayer as expected, but this difference partially levels out during the simulations, due to the rearrangements mentioned above (Figure 3.18). Therefore, coarse-grained representations suggest that the larger deformation observed in the monolayer is due mostly to the decrease in structural robustness when only one layer is present, and only in minor measure to the hydrophobic effect. However the evidence collected through atomistic simulations of assembly of a few capzip molecules suggests the opposite - even if they can tackle a shorter time scale. The atomistic simulation of a monolayer would be a useful piece of information to strengthen the conclusion derived, but was not run in the interest of time.

The discrepancy between coarse-grained and atomistic conclusions can be solved accepting both mechanisms (structural robustness and hydrophobic effect). Indeed, it is clear that coarse-grained descriptions have their weakest point in the ability of reproducing structural solvation, while they provide useful mechanical information.

Based on this, we deemed that the capsules observed experimentally must have a non-monolayer structure: a single layer would not provide enough structural stability and, moreover, it is not compatible with the thickness observed in the TEM images collected experimentally (Figure 1.10(b)).

Overall, the multiscale investigation of this system provided information on how it is structured, but also a useful comparison between force fields, proving how the same system can be described in different ways from different models. This can be useful especially for the (comparatively) new SIRAH force field, which has been less assessed and less widely used than the MARTINI description.



(a) (b)

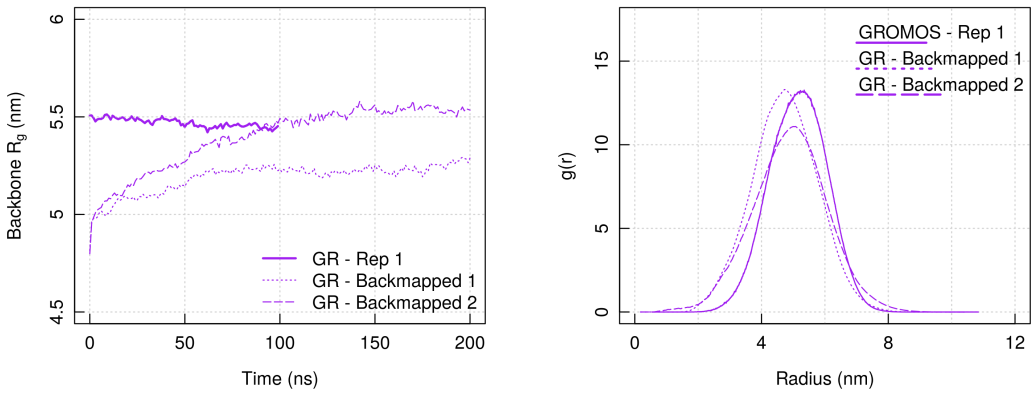
Figure 3.18: Solvent Accessible Surface Area (SASA) per molecule, divided by residue types for simulations of the bilayer and monolayer structure. Results are shown for Replica 1 of each simulation set-up. (a) SASA computed from the initial configuration; (b) from the average over the production run.

3.4.4 Backmapped simulations

Two of the final configurations obtained from standard MARTINI simulations were backmapped to atomistic resolution, and simulated for additional 200 ns. The structures expanded roughly up to the sizes obtained in direct atomistic simulations (Figure 3.19(a)). They present however some unpaired arms and slightly deformed shapes, so that the RDF profiles of their masses around the origin results in a broader curve with respect to the original atomistic simulations, albeit they are centred at the same values (Figure 3.19(b)).

The pattern of contacts per residue, computed on the last 100 ns of simulations, contains features of both the atomistic and MARTINI profiles (Figure 3.20). In particular, at the side chain level, is recovered the interaction between Arginine and Tryptophan residues. Also for mixed backbone-side chain contacts the role of Tryoptophan with respect to the other residues is restored. For backbone contacts instead, the pattern resembles more the one from MARTINI. However, the number of total contacts is slightly lower, probably due to the rapid expansion which follows the backmapping procedure. This movement happens to relieve the unfavourable conformation deriving from the conversion between resolutions.

The expansion of the structure suggests that the atomistic simulations analysed before have indeed reached (or are reaching) an equilibrium: on one side, the atomistic radius of gyration in Figure 3.9(a) has not reached a plateau and could possibly decrease



(a) (b)

Figure 3.19: (a) Radius of gyration (computed on the protein backbone) and (b) RDF of protein masses around their centre for atomistic simulations (Replica 1) and backmapped atomistic from final configurations of Replica 1 and 2 of standard MARTINI runs. The RDF has been averaged over the last 50 ns of the original simulation, or 100 ns of the backmapped simulations. For each label of the legend in (b), the bar has length of the respective FWHM of the Gaussian function fitting the data (thickness estimate).

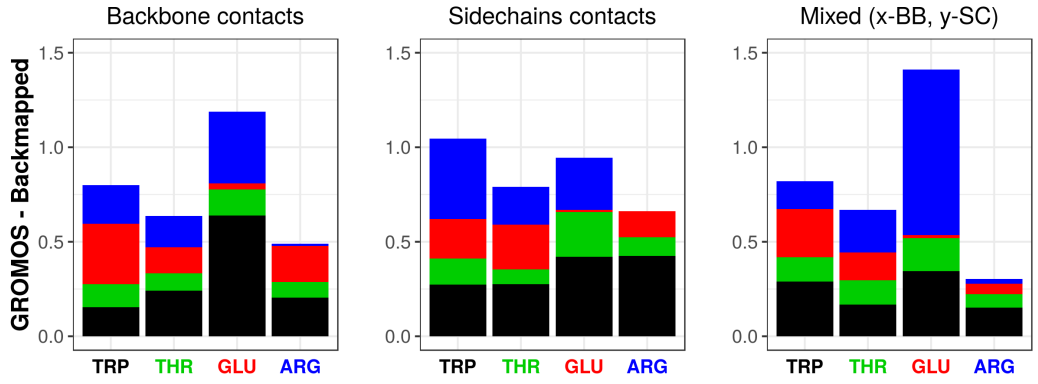


Figure 3.20: Number of contacts per residue type in each arm of capzip for an atomistic simulation backmapped from standard MARTINI (Replica 1): each bar shows the average number for the residue on the x-axis; its color is split by the identity of the partner residue (color coded as in the x-axis legend). For mixed contacts the residue on the x-axis contributes with its backbone. Results are shown for backmapped replica 1. Only contacts existing more than 50% of the simulation time are considered.

more, on the other, when the structure is contracted (as it is after the backmapping) it expands back to the larger size observed in Figure 3.9(a). This suggest that the equilibrium value for an atomistic model lays between what found with an atomistic and a MARTINI coarse grain approach.

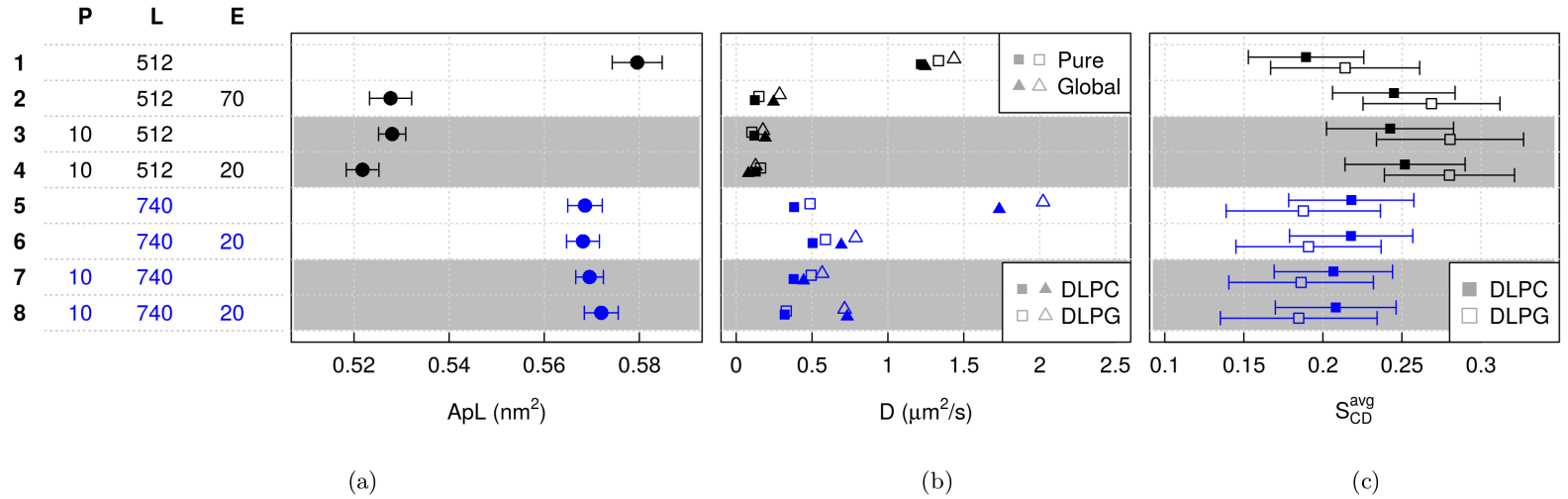
3.5 Results: peptide-membrane interactions

We now discuss simulations of capzip in contact with a membrane. First we focus on atomistic simulations of the pentagonal peptidic subunit in contact with a model bacterial or mammalian membrane, both under standard simulations conditions and with an applied external electric field. In this, we elucidate the local effect that the peptide has on the local lipid organisation. We then analyse the process of membrane binding which is observed in coarse-grained simulations, comparing the MARTINI and Polar MARTINI force fields for the simulation without an external electric field, and necessarily resorting to Polar MARTINI for the ones involving it.

3.5.1 Atomistic simulations of the bacterial model membrane

Preliminary simulations were run on bacterial membranes alone (two lipid bilayers with 512 and 740 lipids respectively), to equilibrate them and compute their characteristics in absence of the peptide. For the bacterial model chosen (DLPC/DLPG 3:1), we obtained an ApL of 0.580(5) nm^2 for the 512-lipid bilayer (simulated with GROMOS 54A7 force field), and 0.569(4) nm^2 for the 740 one (GROMOS 54A8) as shown in Figure 3.21(a), lines 1 and 5. No experimental values are available for this lipid mixture to compare the computational results.

The difference is small but statistically significant (the two measures are not compatible within the error). There are three differences between the two sets of simulations, and each of them can be responsible for the discrepancy. First, the force field (GROMOS 54A7 versus 54A8, see Chapter 4 for a thorough comparison), then the long range electrostatic treatment (RF versus PME), and finally the size of the system (512 versus 740 lipids). Regarding the first, extensive tests on phosphocholine lipids (see Chapter 4) suggest that generally 54A8 gives larger ApL values than the ones computed with 54A7. Regarding the second, a control simulation on the 512-lipid bilayer was run swapping the RF long range electrostatic treatment with PME (and maintaining the remaining set-up). It produced an ApL of 0.586(5), slightly larger that the one obtained with RF, but compatible with its value. Therefore we attribute the discrepancy to the different sizes of the systems. Later in this section other cases are



(a) (b) (c)

Figure 3.21: (a) Area per lipid (ApL); (b) pure and global lateral lipid diffusion coefficient (D), (c) average tail order S_{CD}^{avg} from simulations of model bacterial membrane bilayers under different conditions. On the left of (a) a schematic of the system set-up shows the number of peptides P, number of lipids L, and electric field E applied in the z direction (in mV/nm) in each case. Black points refer to simulations on 512-lipid bilayers, blue ones on 740-lipid bilayers; shadowed regions flag systems where the peptide is present; in (b) squares refers to pure diffusion, triangles to global diffusion; in (b) and (c) full symbol are for DLPC, hollow ones for DLPG. In (b) and (c) staggering along the y-axis helps the visualisation. Bars denote the standard error; for the diffusion coefficients, these are smaller than the point size. Analysis performed discarding the first 200 ns of the simulations time.

reported in which the size has a great influence on the simulation outcome.

Regarding the diffusion coefficient, we report in Figure 3.21(b) both the pure and global values (computed as explained in Section 3.3 and Figure 3.7). Considering the pure diffusion (squares in Figure 3.21(b)), the larger membrane has a smaller coefficient with respect to the 512-lipid bilayer (line 1 and 5 of the Figure above). However, considering the global one (triangles) we observe the opposite. In general, differences are expected between the two systems, because of their different size: it is well known that periodic boundary conditions affect the computed diffusion coefficient D [Camley et al., 2015; Venable et al., 2017]. Previous studies found that, for atomistic simulations of the size employed here, D is systematically underestimated [Camley et al., 2015]: increasing the size would increase the coefficient, toward its real value (which is obtained for an infinite box). This is consistent with our global diffusion results, but not the pure ones. Unfortunately, it is not clear whether, in the publications mentioned, they refer to what is here defined as pure or global diffusion.

To exclude that the difference between the two systems is given by the force field chosen, we refer the reader to the next chapter, where, for simulations of monolipid bilayers, the 54A7 or 54A8 force fields gave very similar values of pure diffusion, keeping all the other simulations conditions identical (of the order of 0.1 $\mu m^2/s$). Also in those simulations we observed drift of one leaflet with respect to the other, which would produce discrepancy between the values of pure and global D. The drift appeared in some of the simulations only, and not in a consistent manner: it did not happen for a lipid in particular, nor for a specific set of parameters. This suggests that small changes in the the initial conditions can lead to very different values of the drift (if any exist). However, once it is started, such motion seems to persist throughout the simulation length, consistently with the inertia gained by the leaflet.

Interaction of the peptide with the 512-lipid model bacterial membrane

For the small bilayer, the presence of the pentagonal peptide subunit (made of ten antimicrobial molecules, and bearing a $+60\,e$ total charge) made the ApL decrease by 9% with respect to the value of the membrane itself (Figure 3.21(a), line 3 versus 1). A similar effect was observed applying an electric field of 70 mV/nm intensity or

combining the peptide with a 20 mV/nm field, comparable with physiological values (Figure 3.21(a), lines 2 and 4). Given the high net charge of capzip, the molecule itself produces an electric field, thus the perturbation on the membrane is similar to the one obtained when the field is externally applied.

The fact that both the peptide and/or the electric field lead to reduction of ApL of a similar amount suggests that the membrane is approaching the maximum packing allowed. Consistently, the diffusion coefficients had a 6-fold reduction for both species of lipids (Figure 3.21(b)). Pure and global diffusion are very similar, as there is little translation between the leaflets. At the same time, the order parameter S_{CD}^{avg} increased slightly (Figure 3.21(c)). The variability on S_{CD}^{avg} is high, showing that some regions reached a value larger than 0.3, which is often associated to transition to the gel phase, characterised by highly order lipid chains [Pluhackova et al., 2016]. To investigate this, for the simulation in presence of capzip, the hexagonal order parameter S_6 is computed for the bottom leaflet of the membrane (the one not hosting the pentagonal subunit, if present). This shows a few lipids in the gel phase, i.e. $S_6 \geq 0.72$ (Figure 3.22): despite there is no formation of ordered clusters, the shrinking of the membrane due to the presence of capzip increases the lateral order (and the thickness of the membrane). Indeed no lipid reaches the S_6 gel threshold in the simulations of a pure membrane (data not shown).

Interaction of the peptide with the 740-lipid model bacterial membrane

When we simulated the 740-lipid bilayer with an externally applied electric field and/or in presence of the peptide, the ApL and tail order parameter S_{CD}^{avg} remained equal, within the error, to the values obtained from simulations of the membrane alone.

The global diffusion coefficient D was still significantly affected, being reduced between 2.5 and 5 times (Figure 3.21, lines 5 to 8), but not the pure diffusion. The simulation of the 740-lipid bacterial membrane alone is the only case, among all the simulations results shown in Figure 3.21(b), in which pure and global D differ substantially. However, as commented above, the leaflet translation seems to be dependent on the initial conditions, thus we can not be certain that this translation is a consistently reproducible effect.

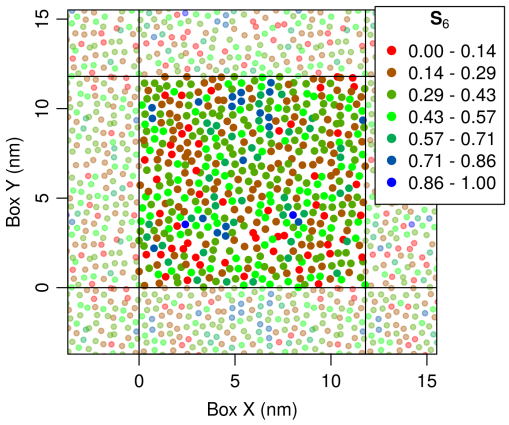


Figure 3.22: Hexagonal order parameter S_6 for the bottom leaflet lipid acyl chains, computed on the last frame of a simulation of a 512-lipid bacterial bilayer with 10 peptide molecules (line 3 in 3.21). Chains plotted by the average xy position of their carbon atoms, and colour coded by the S_6 value; chains of the periodic images shown faded out; boundaries of the simulation box in solid black lines.

While, for the small bilayer, perturbations reduced the diffusion and the ApL together, the results on the larger membrane show that neither ApL nor pure diffusion are affected from the peptide/electric field. However, given the decrease in global diffusion, an the fact that the peptide is likely to influence the movement of the lipids nearby itself, we computed the diffusion coefficient for subsets of lipids based on their distance from the peptide at the initial frame of the production run (Table 3.2). On this occasion, the trajectory was centred around the peptide COM, to understand the lipid movement with respect to it. The lipids closer to the protein resulted indeed the most slowed down in their motion. DLPG is slightly more mobile than DLPC, as observed in all the 740-lipid bilayer simulations (Figure 3.21(b)).

The local slowing down of lipids is due to electrostatic and hydrogen bond interactions between the lipids and the peptide. We computed the latter, monitoring the ones which persisted more than 50% of the time, not necessarily i n a continuous fashion (Table 3.3). Moreover, we singled out the ones involving Arginine and Tryptophan residues. Arginine promotes many persistent bonds, and the number of them is equally divided between DLPC and DLPG residues, despite there are less DLPG residues overall. Tryptophan promotes fewer persistent bonds; and the remaining ones that exist for more than 50% of the time involve all Threonine residues. Thus, hydrogen bonds, together with the electrostatic attraction between positive residues of the peptide and

Local diffusion for DLPC/DLPG with capzip

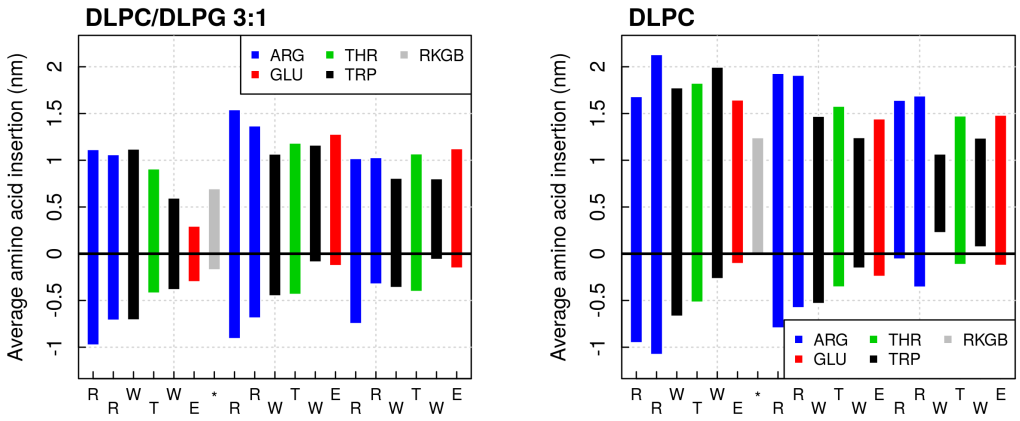
			DLPC		DLPG	
Lip.	Pept.	Region	Nr.	D ($\mu m^2/s$)	Nr.	D ($\mu m^2/s$)
10	737	$d < 1$	93	0.283(1)	38	0.272(2)
10	737	$d < 2$	152	0.325(2)	52	0.353(3)
10	737	$d < 3$	239	0.402(2)	70	0.393(3)
10	737	$d > 3$	329	0.442(3)	98	0.650(4)
10	737	All	568	0.486(3)	169	0.541(3)
0	740	All (global)	570	1.75(3)	170	2.02(2)
0	740	All (pure)	570	0.383(1)	170	0.487(3)

Table 3.2: Local diffusion for the DLPC/DLPC bilayer in presence of capzip. Pept.: number of peptides; Lip.: number of lipids. The values are computed for groups of lipids which, at the initial time, are at a distance d smaller than 1 nm, 2 nm or 3 nm from the peptide or larger than 3 nm (Regions). D is computed centring the trajectory around the Protein COM. The pure and global diffusion coefficients for the pure membrane are given for comparison. Error from linear fit in parenthesis.

Capzip - DLPC/DLPG lipids hydrogen bonds

	E = 0 mV/nm				**E = 20 mV/nm**			
	Total		$\tau \geq 50\%$		Total		$\tau \geq 50\%$	
	R	W	R	W	R	W	R	W
DLPC	668	73	20	7	798	90	18	6
DLPG	928	114	12	3	980	104	16	3
All	2041		50		2238		54	

Table 3.3: Number of hydrogen bonds between the Arginine (R) or Tryptophan (W) residues of the pentagonal subunit (10 capzip molecules) and DLPC or DLPG lipids (740-lipid bacterial bilayer). Results shown for simulations without electric field (left) or a 20 mV/nm field (right). The table displays both the total number of observed hydrogen bonds (Total) and the ones present for more than 50% of the simulated time, not necessarily in a continuous fashion ($\tau \geq 50$). Furthermore, the corresponding figures for all the bonds occurring between any residue of the peptide and any lipid is displayed (All).



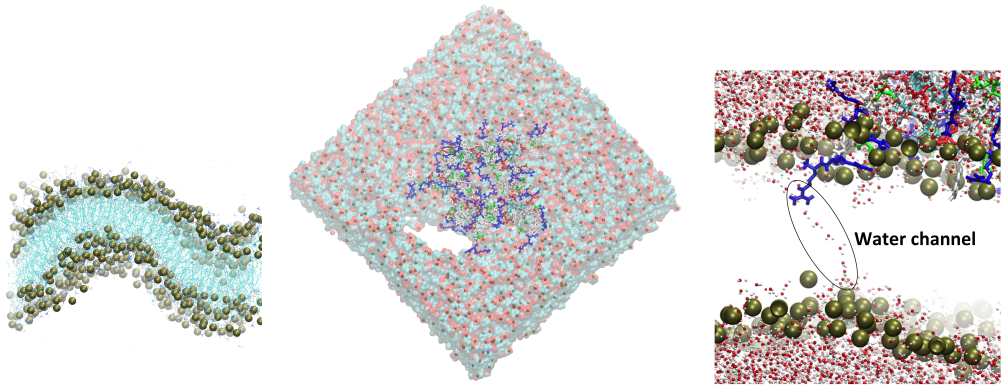
(a) (b)

Figure 3.23: Insertion of each amino acid of the pentagonal subunit within a bacterial (a) or mammal (b) model membrane. The x-axis shows the amino acids in the capzip sequence, with · indicating the RKGB central residue (19 overall). The boxes shows the range on the 10 molecules forming the pentagonal subunit.

negatively charged DLPG lipids, slow down the dynamics of the lipids close to the peptide, which do not diffuse away. The presence of a less mobile patch of lipids influences the diffusion of the ones nearby as well, decreasing it.

Regarding the position of the amino acid within the lipid membrane, Figure 3.23(a) shows a plot of the average insertion during the last 100 ns of the simulation. This is computed taking into account for each amino acid the atom with the lowest z coordinate, which is compared with the maximum z coordinates of the lipids around it. In this way, the local height of the membrane is taken into account. The results show a deeper insertion of Arginine residues on average, while Tryptophan inserts less, and some of them are placed at the interface with solution. This will be important for the initiation of pores in simulations with an external electric field (see the next paragraph).

We think that this reduced "fluidity" in terms of lateral diffusion is an important perturbation of the membrane structure, as it diminishes its ability to accommodate external stimuli, such as water penetration. With coarse-grained simulations (see Section 3.5.3) it will be evident that on the long time scale, these interactions are also able to recruit anionic lipids around the peptide, creating a local imbalance in the membrane composition.



(a) (b) (c)

Figure 3.24: (a) Bacterial membrane deformation due to an electric field applied to the membrane (512 lipids, E = 130 mV/nm). (b) Pore formation due to the action of the peptide and electric field of 130 mV/nm. (c) Pore precursor due to Arginine insertion and water penetration. [VMD software, Humphrey et al. [1996]]

Electroporation results As it is not possible to witness the penetration of the peptide through the membrane within the available simulation time, we opted to perform electroporation simulations. An electric field of increasing intensity was applied along the negative z direction perpendicular to the membrane, with the peptide on the positive z side, to model the field generated by the transmembrane potential.

The field was increased of 20 mV/nm every 200 ns (or 10 mV/nm when approaching the poration threshold). The initial value of 20 mV/nm was chosen as it is an approximation of the physiological value of the transmembrane potential (see discussion in Section 3.2). This procedure showed that the critical value of 130 mV/nm triggered poration in the presence of the peptide. This was confirmed by three replicas run with a 130 mV/nm field from the initial unperturbed membrane configuration (poration after 20 ns, 75 ns and 71 ns respectively). Similarly, on the 740-lipid bilayer in presence of capzip, the same field induced membrane disruption after 60 ns, 50 ns and 70 ns respectively (Figure 3.24(b)). This value is significant at the time scale used, and it is possible that for longer time scales one can witness poration at lower values of the electric field.

As a control, a pure 512-lipid bacterial membrane was simulated under the same conditions, in three replicas: in the 600 ns runs, we observed the appearance of curved regions (Figure 3.24(a)) but no poration. The appearance of a curvature made it

necessary to compute the area per lipid taking this into account, as explained in Section 3.3. The three replicas gave values of 0.520 nm^2, 0.514 nm^2 and 0.550 nm^2. Their discrepancy is due to the different level of curvature the membrane adopts during the runs (as they can change at a fast pace, the values given above were computed over the last 10 ns of the trajectory only). It is observed that once a curvature appears it can be quickly enhanced by the electric field. Small casual variations in its initial onset can bring to very different shapes of the membrane and thus of compression of the lipids in it, giving different ApL values.

The electroporation threshold for the bacterial membrane alone was set at the higher value of 140 mV/nm: out of three simulations run with such value of the electric field, two resulted in disruption at 150 ns and 154 ns, while the third presented a curved but still intact membrane after 200 ns. This shows that the effect of the field combined with the presence of the peptide accelerates the disruption process. Because of that, a slightly lower field value was sufficient to observe poration when the peptide was present.

The simulations with high electric field performed on the 512-lipid bilayer resulted all in curved geometries before the poration event (independently from the presence of the peptide). However this does not happen with the larger membrane, for which the poration was always initiated from a flat conformation (Figure 3.24(b)). To rule out once more the hypothesis that the difference is due to the force field used (GROMOS 54A7 versus 54A8), we run an additional control simulation of the 512-lipid bilayer with the 54A8 force field and the electric field set at the electroporation threshold of 140 mV/nm. The bilayer, which was not electroporated in the 200 ns run, developed a curved shape similar to the outcome from the corresponding 54A7 simulations. Also, the ApL values of the two simulations of the 512-lipid bilayer with peptide and a 140 mV/nm electric field, differing only for the force field (54A7 or 54A8), are quite similar: 0.554 versus 0.556 nm^2, computed taking into account the curvature. However, from the previous discussion it emerges that there is a high variability in the ApL computed from highly curved configurations, so that the consistency of the two values might be fortuitous. In any case, this analysis proves that the size of the bilayer has a large influence in the outcome of the simulations.

Arg Capzip - DLPC/DLPG lipids hydrogen bonds

$[\tau \geq 50\%]$	**E = 0 mV/nm**			**E = 20 mV/nm**		
	Total	Pho	CO_n	All	Pho	CO_n
DLPC	20	6	14	18	10	8
DLPG	12	3	9	16	5	10

Table 3.4: Number of the hydrogen bonds between Arginine residues and lipids, existing for more than 50% of the time, in simulations of the 740-lipid bacterial bilayer with the peptide and electric field equal to zero or 20 mV/nm. Total number is displayed and the ones occurring with the Phosphate (Pho) moiety of the lipids or one of the ester groups (CO_n).

Regardless the shape that the membrane assumes before disruption, the peptide speeds up the membrane disruption. This is due to the charged Arginine residues which insert into the membrane core interacting with the negatively charged phosphate group of the lipids. They form hydrogen bonds with the ester groups, promoting the penetration of water molecule (Figure 3.24(c) and Supplementary Movies SI_M4). The precursor of this mechanism can be seen also in the simulations without or with low electric field (20 mV/nm): indeed, the vast majority of hydrogen bonds between Arginine and lipids listed in Table 3.3 involves the oxygen of the ester or phosphate group of the lipids, as shown in Table 3.4. This is also consistent with the deeper insertion observed for Arginine residues with respect to other ones (Figure 3.23(a)).

The positive charge of Arginine residues makes its behaviour within the membrane complex: Arginine residues buried in the membrane, can either bend their chain to re-surface in solution (in a process called snorkelling [Liang et al., 2005; Ulmschneider et al., 2017; Öjemalm et al., 2016]), or make interaction with the phosphate and carbonyl groups, as observed in our simulations. The second option has been observed for Arginine rich antimicrobial or cell penetrating peptides through NMR experiments which measured the distance between the nitrogen of Arginine and the Phosphorus of lipids [Tang et al., 2007; Jobin et al., 2019] and it was confirmed by MD simulations of the same or similar systems [Herce et al., 2009; Jobin et al., 2019], consistently with what found in this work.

Finally, capzip causes a slight deformation on the shape of the membrane. This, together with the Arginine insertion, allows pore formation at a value of the electric

field lower than the one necessary for electroporation.

Therefore, the presence of capzip has two main effects: it decreases the membrane fluidity (as measured by the reduction of the diffusion coefficient around the peptide), and makes the bilayer more sensitive to electric field, favouring the formation of water channels - two effects which are likely correlated, and which contribute to trigger its membrane disrupting activity.

3.5.2 Atomistic simulations of the mammalian model membrane

After the investigation of a model bacterial membrane, we focussed on a mammalian one, modelled as a DLPC membrane. We opted for a pure DLPC membrane to have a mammaloan model as close as possible to the bacterial one (DLPC/DLPG), to reduce the number of variables while testing the action of the peptide on these membranes. Thus, the mammalian counterpart was obtained selecting only the zwitterionic lipid species of the bacterial one. In turn, the bacterial model composition was selected to be the same as in the Supported Lipid Bilayers experiments [Castelletto et al., 2016]. Given the information accumulated in the previous investigation, we simulated directly a large bilayer (748 lipids), as we deemed this size more appropriate to run the subsequent simulations with the peptide. The force field employed is GROMOS 54A8 [Reif et al., 2012].

The properties of the stand alone membrane are computed as before and compared with the experimental data. Indeed for mono-lipid membranes several experimental results are available. The computed ApL is 0.592(3) nm^2, while the experimental range at 303 K of temperature is 0.608-0.632 nm^2 (see Table 1 in Poger et al. [2016]). To motivate the discrepancies, two factors must be considered. First, these experiments are performed on a variety of lipid geometries, from vesicle to Supported Lipid Bilayers (resting on a solid surface), justifying the broad range of values. Second, the membrane is simulated in presence of a 150 mM salt concentration, while the experiments do not adopt it. Several computational studies found a reduction of area per lipid when salt concentration is introduced [Böckmann et al., 2003; Jarerattanachat et al., 2013; Reif et al., 2017]. This is consistent with the experimental evidence [Pabst et al., 2007], however, Reif et al. [2017] reported that simulations often overestimate this variation.

Properties of DLPC - atomistic

Pept.	Lip.	C (mM)	ApL (nm^2)	D (μm^2/s)
–	748	150	0.592(3)	0.673(1) (global)
–	748	150	0.592(3)	0.495(1) (pure)
–	512 [a]	0	0.626(5)	0.541(1) (pure)
–	Exp.[b]	0	0.608-0.632	3
10	748	150	0.592(4)	see Table 3.6

Table 3.5: Area per lipid and diffusion coefficient (global or pure) from atomistic simulations of a DLPC bilayers. Pept.: number of peptides; Lip.: number of lipids. [a] Run with RF long range electrostatics and 1.4 nm cut off. [b] Experimental values: area per lipid from Table 1 of Poger et al. [2016], diffusion coefficients from Lindblom and Orädd [2009].

The fact that salt is responsible for the reduction in ApL is supported by the results from a simulation of a 512-lipid DLPC bilayer without salt, which gives ApL of 0.626(5) nm^2, compatible with the experimental values (see Table 3.5 for a summary. To be noticed though that we are comparing two different sizes). Analogously, simulations of the 512-lipid bacterial model membrane without salt gave ApL of 0.596(5) nm^2, which is 3% higher than the values found with a 150 mM concentration of NaCl (0.579(5) nm^2). Moreover simulations performed without salt (see Chapter 4 show a better agreement with the experimental values. Therefore, we conclude that the effect of the salt concentration is responsible for the smaller ApL observed for the large DLPC membrane simulated with 150 mM NaCl with respect to the experimental values.

The ApL of the DLPC bilayer is larger than the one found for the model bacterial membrane, likely because the presence of DLPG in the latter diminishes the repulsion between the positively charged Choline heads of DLPC. Accordingly, the pure lateral diffusion (0.495(1) μm^2/s) is slightly higher than with what found for the mixed membrane (0.383(1) μm^2/s for DLPC and 0.487(3) μm^2/s for DLPC). Simulations of a 512-lipid DLPC bilayer without salt (see Figure 8 in Chapter 4) gave a higher diffusion value with respect to the 150 mM NaCl simulation, and again we deemed the presence of the salt responsible for this.

The presence of the peptide does not affect the ApL of the DLPC bilayer, as it was observed for the bacterial counterpart (on the same system size). Computing

Diffusion for DLPC with capzip

Pept.	Lip.	Region	Nr.	D ($\mu m^2/s$)
10	748	$d < 1$	100	0.495(3)
10	748	$d < 2$	180	0.584(3)
10	748	$d < 3$	323	0.603(5)
10	748	$d > 3$	412	0.993(5)
10	748	All	735	0.821(2)
0	748	All (B_{COM})	748	0.673(1)
0	748	All (L_{COM})	748	0.495(1)

Table 3.6: Diffusion coefficients of lipids, in a mammalian model membrane. Pept.: number of peptides; Lip.: number of lipids; Region: the values are computed for groups of lipids which, at the initial time, were within 1 nm, 2 nm, 3 nm from the peptide or further away. D is computed centring the trajectory around the Protein COM, when this is present. The pure and global coefficients are given for the pure membrane. Error from linear fit in parenthesis.

the diffusion centring the trajectory with respect to the peptide, we observe a larger diffusion coefficient with respect to the one computed for the membrane alone, likely because of the movement of one leaflet with respect to the other. Moreover, the lipids closer to the peptide are still slowed down in their movement as observed in the bacterial membrane (Table 3.6).

Analysing the number of hydrogen bonds formed with the peptide, there are 590 bonds, of which 19 present more than 50% of the time. Of them, 11 involve Arginine residues coordinated with ester or phosphate groups. These figures are smaller than the ones obtained for the bacterial membrane. In particular, the absence of the charged lipids makes the protein-membrane interaction less favourable. This is confirmed by the measure of the amino acid insertion (Figure 3.23(b)) which are on average less inserted than into the bacterial membrane.

The investigation performed on the bacterial model membrane sets to 130 mV/nm the threshold electric field needed to observe poration at the atomistic level on such membrane (with the GROMOS force field). Therefore, we want to reproduce this condition on a DLPC model membrane and see whether the influence of capzip is smaller on the DLPC membrane with respect to the DLPC/DLPG. However, for the

DLPC membrane this value was sufficient to cause poration both with and without peptide. Despite the absence of lipids with a net charge, which are expected to interact more strongly with the electric field, this membrane seems more sensitive to such external perturbation. We think this is due to larger ApL: indeed, being the lipids less packed, water penetration is easier. This, combined with the electric field, generates more easily water channels which promote pore formation.

In the case of the mammalian model membrane, we observed that, once the membrane starts disrupting, the peptide detaches from it, while it remained firmly attached in the simulations with the bacterial one, following its deformations. This lower propensity for binding on the mammalian case will be proven also by coarse-grained simulations and is consistent with the previous analysis on hydrogen bonds propensity and amino acids insertion. Moreover it confirms that the poration is induced because of the high electric field rather than the presence of the peptide. We do not pursue the investigation of the electroporation threshold for the DLPC membrane, as we focussed here only on the action of the peptide on it. Finally, in interpreting these results, it must be considered that the transmembrane potential of a mammal cell is lower than the one present in the bacterial membrane [Yeaman and Yount, 2003; Wilson et al., 2011], so that their membranes usually need to withstand milder electric field conditions than bacterial ones.

3.5.3 CG simulations of the buckyball on model membranes

Standard MARTINI simulations Coarse-grained simulations allow to model the behaviour of the full capsule interacting with the membrane. At first, simulations with the standard MARTINI model were run, as they have higher computational speed with respect to the Polar MARTINI model. To be noticed that the full system, comprising a 2880 lipids bilayer and the capsule, measures approximately 30 nm along each side of the simulation box.

The coarse-grained simulations of the full buckyball bilayer on a DLPC/DLPG membrane (3:1 ratio) confirms the binding of the peptide on the latter, driven by charge-charge recognition: in both the replicas run, the peptide approaches the membrane after about 2 μs, remaining bound for the remaining of the 10 μs simulated. Post-binding,

the capsule diffuses on the membrane and produces an increasingly high curvature on it, in a process which tends to maximize the contact area (see Figure 3.5(b) and Supplementary Movies SI_M5). No poration is observed, probably due to the force field characteristics which stabilise the structure of both the membrane and the peptide assembly. Additionally, longer time scales might be needed to observe it. Later in this section, we will analyse simulations performed with an externally applied electric field, in line with what done in the atomistic case, which speeds up the process and allow to observe membrane disruption.

The area per lipid was computed taking into account the curvature (as explained in Section 3.3 and in Braun et al. [2011]), and does not change significantly in presence of the capsule. When analysing the diffusion coefficient, the values obtained for pure membranes (or before the binding) have been computed removing the bilayer centre of mass movement, while the ones with the capsule bound to the membrane by centring the trajectory around it (and considering the portion of the trajectory after the binding only). The MARTINI diffusion coefficients are about 60 times larger than the ones obtained atomistically (Table 3.7), due to the coarse-grained parametrisation. The values are consistent with results obtained on a variety of other systems tested by Venable et al. [2017] in a study of diffusion for MARTINI simulations.

The diffusion coefficients obtained for simulations of the pure membrane and for simulations with the buckyball bilayer before its binding (lines 1 and 5 in Table 3.7, top) are similar for the DLPC/DLPG membrane and shows that the two lipids diffuse at the same pace. The lipids diffusion is instead decreased after binding of the buckyball bilayer, by 5% for DLPC and by 15% for DLPG (line 6 of Table 3.7, top). The larger effect on DLPG is due to the attraction between the peptide and this negatively charged lipid: computing the RDF of the phosphate beads (PO4) of DLPC and DLPG around the protein, a much stronger signal comes from DLPG beads rather than DLPC ones (Figure 3.25(a)), showing they are more attracted toward the peptide.

The mechanism above results in the slower diffusion of the DLPG residues nearby the capsule, as already predicted at the atomistic level. Moreover, the coarse-grained simulations allow to see the recruitment of negatively charged lipids nearby the capsule: Figure 3.25(b) shows a lateral density profile of the two lipid species across the

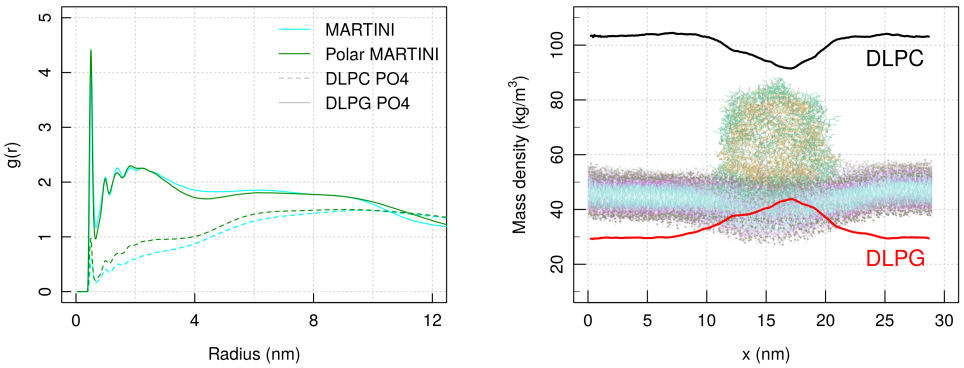
Properties of DLPC/DLPG - MARTINI

	Pept.	Lip.	C (mM)	E (mV/nm)	FF	ApL (nm^2)	D_{DLPC} ($\mu m^2/s$)	D_{DLPG} ($\mu m^2/s$)
1	–	2880	0	–	MA	0.581(3)	75.84(7)	76.64(5)
2	–	2880	150	–	MA_P	0.611(1)	79.4(4)	79.0(5)
3	–	2880	150	20	MA_P	0.614(1)	74.64(5)	73.67(2)
4	–	2880	150	40	MA_P	0.619(2)	70.31(4)	70.45(6)
5	120 (BB)^b	2880	0	–	MA	0.582(1)	74.28(1)	73.34(2)
6	120 (BB)^a	2880	0	–	MA	0.581(2)	70.51(1)	62.59(2)
7	120 (BB)^a	2880	150	–	MA_P	0.621(2)	65.82(6)	68.94(9)
8	120 (BB)^a	2880	150	20	MA_P	0.616(2)	70.29(7)	61.54(10)
9	120 (BB)^a	2880	150	40	MA_P		Poration	

Properties of DLPC - MARTINI

	Pept.	Lip.	C (mM)	E (mV/nm)	FF	ApL (nm^2)	D_{DLPC} ($\mu m^2/s$)
1	–	2888	0	–	MA	0.590(2)	83.5(3)
2	–	2888	150	–	MA_P	0.608(2)	80.5(4)
3	–	2888	150	20	MA_P	0.610(1)	78.3(1)
4	–	2888	150	40	MA_P	0.614(1)	76.90(3)
5	120 (BB)	2888	0	–	MA	0.590(1)	81.07(3)
6	120 (BB)	2888	150	–	MA_P	0.608(1)	70.52(2)
7	120 (BB)	2888	150	20	MA_P	0.611(1)	75.28(6)
8	120 (BB)	2888	150	40	MA_P	0.616(1)	77.88(4)

Table 3.7: Properties of DLPC/DLPG (top) and DLPC membrane (bottom) in MARTINI simulations with and without capzip. Pept.: number of peptides (BB - buckyball bilayer). Lip.: number of lipids; FF: force field (MA - MARTINI; MA_P - Polar MARTINI). D is computed removing the centre of mass movement of the bilayer for simulations without the capsule, and movement of the protein for simulations where the peptide is bound to the membrane. For DLPC/DLPG, the buckyball bilayer binds to the membrane: values marked with ^{b} are computed *before its binding*, with ^{a} *after the binding*. For DLPC it never binds and the values are computed *for the whole simulation length* (discarding the first 100 ns).



(a) (b)

Figure 3.25: (a) RDF of phosphate bead (named PO4 in both MARTINI models) of DLPC and DLPG around the Protein. (b) Mass density of DLPC and DLPC in a DLPC/DLPG membrane along the x-axis, computed for simulations of the membrane with the buckyball bound to it. Superimposed is a picture of the corresponding configuration.

simulation box. Spanning the x axis, the DLPG density increases around the position of the capsule, at the expenses of the DLPC one. The formation of a DLPG patch enhances the slowing effect on the diffusion of this species, while DLPC, pushed away from the capsule location, is less perturbed.

To conclude the investigation performed with standard MARTINI, we simulated the capsule together with a pure DLPC membrane. The capsule does not bind to the latter in the 10 μs simulated (Supplementary Movies SI_M6): it comes close to it multiple times, keeping an average distance of 3 nm and a minimum of 1 nm (Supplementary Figure 3.41). This, together with what already observed in the atomistic simulations, suggests once more that the interaction with non charged lipids is not leading to attachment of the buckyball to the membrane.

Polar MARTINI simulations Polar MARTINI simulations focussed on *in silico* experiments with an external electric field, to set a parallel with the analogous investigation performed at the atomistic level. The addition of an external electric field is not possible in standard MARTINI simulations because standard water does not provide a long range electrostatic screening. As such, the electric field effect would be overestimated. Together with the simulations of the membrane with the capsule, we run control simulations of the pure membrane with the selected values of the external field,

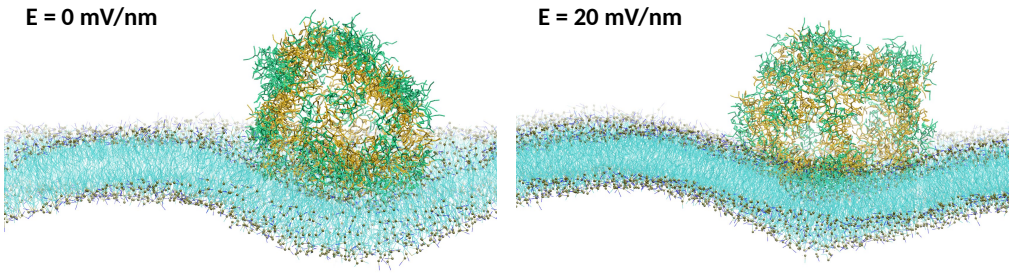
to exclude the possibility of electroporation.

Simulations on the membranes alone (both the bacterial and mammalian model) showed an increased ApL with respect to the standard MARTINI simulations (lines 1-2 in Tables 3.7, top and bottom). This is the opposite of what observed in the test case run in the original work that parametrised Polar water [Yesylevskyy et al., 2010], possibly due to the different size and composition of the test membrane used. Additionally, the two simulations are run with different values of the salt concentration (MARTINI has only counter ions, Polar MARTINI has 150 mM NaCl): in the atomistic case this discrepancy made the ApL decrease, which is again inconsistent with what observed here and can be attributed to the parametrisation of ions in the coarse-grained model (e.g. their van der Waals radius includes the first hydration shell, thus they will penetrate less easily into the lipid head region). Regarding the diffusion coefficient it decreases from MARTINI to Polar MARTINI, despite an increase in ApL, and this is likely due to stronger interactions between lipids (especially the charged DLPG). This is due to the Coulomb interaction, which is less screened and thus stronger in the Polar MARTINI model.

When an electric field is applied in Polar MARTINI simulations, there is a slight increase of ApL for higher values of it, but the diffusion decreases slightly (lines 2-4 in Tables 3.7, top and bottom). In atomistic simulations the field decreased the ApL for small systems and had negligible effect for larger ones; while the diffusion was reduced, as in the Polar MARTINI case.

To simulate the binding of the buckyball bilayer, we first focussed on the bacterial membrane. We run one simulation without electric field, to test the interaction between the two with the new parametrisation. The binding happens much faster than what observed in the standard MARTINI simulations (within the first 50 ns), and the membrane starts to be deformed by the capsule already after 100 ns (Figure 3.26(a)). For this reason, Table 3.7 does not report a value of the diffusion *before* binding.

We did not observe poration on the time scale of the simulation (500 ns): a longer simulation might allow to observe membrane disruption but, as mentioned before, we preferred to focus on the action of the electric field. As such, we simulated the binding under the action of a 20 mV/nm electric field. The curvature of the membrane is slightly



(a) (b)

Figure 3.26: Final configuration (500 ns) of Polar MARTINI simulations of a buckyball bilayer approaching and binding to a model bacterial membrane, spontaneously (a) or under the effect of a 20 mV/nm electric field (b). Capsule in bond representation, green external layer, yellow internal one. Lipids in line representation, phosphate beads in gold van der Waals. [VMD software Humphrey et al. [1996]]

more pronounced than in the case without field, but again no disruption happens in the 500 ns simulated (Figure 3.26(b)). From the final configuration of this run we doubled the electric field, which was sufficient to observe poration within 200 ns. Remarkably, a value of 40 mV/nm is within the physiological range of the transmembrane potential (which is around 35 mV/nm for the bacterial inner membrane, and 20 mV/nm across a mammalian membrane. See discussion in Section 3.2).

As in membrane simulations the pressure coupling is performed semi isotropically, once a pore is formed, the box undergoes a large and unphysical deformation if the pore keeps expanding. The analysis of the events happening after poration are thus affected by this artefact. To obviate to that, we took a frame at the early stages of pore formation (1 ns after the first formation of a water channel) and continued the run, with the same electric field, with isotropic pressure coupling. This allowed to observe the capsule penetrating the membrane in the first 10 ns from the beginning of this run (Figure 3.27 and Supplementary Movies SI_M7): the lipids do not seal around the capsule, allowing the passage of water and ions, which is consistent with what found in the atomistic simulations. Moreover, in the passage, the capsule deforms and partially opens. After this, the membrane is completely disrupted and the lipids rearranged into fibers ad micelles to shield their tails from the solvent, while some of them remain attached to the capsule (Supplementary Movie SI_M7).

The fact that we observed poration at a much lower value of the electric field than

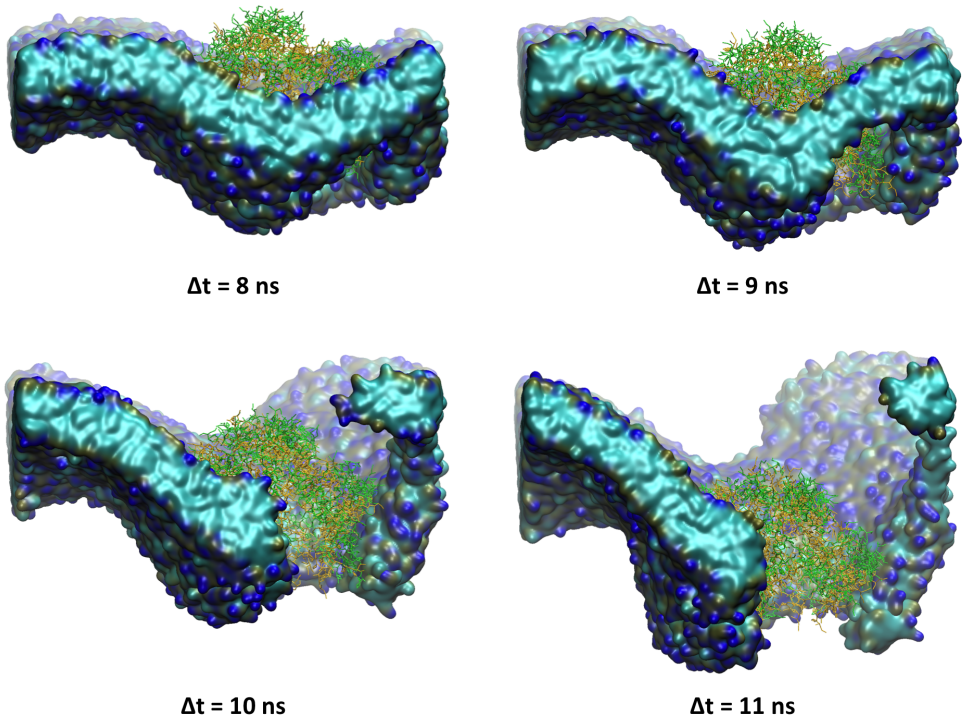


Figure 3.27: Bacterial membrane poration due to the action of the capzip buckyball bilayer and an external electric field of 40 mV/nm (Polar MARTINI simulations). Δt indicates the time from the beginning of the simulation performed with isotropic pressure coupling (started 1 ns after the formation of a water channel spanning the membrane). Capsule in bond representation, green external layer, yellow internal one. Lipids in surface representation. [VMD software Humphrey et al. [1996]]

what found with the atomistic model can be explained by several characteristic of the coarse-grained parametrisation: first, it favours a more dynamical rearrangement between lipids (see the discussion on diffusion in the following); second, it is known for smoothing the energy barriers and thus speeding up the course of events: the effective time sampled in MARTINI simulations is 3 to 6 times larger than the one actually simulated [Marrink et al., 2004]. Moreover, in these simulations the full capsule with 120 molecules is present, interacting with the electric field more strongly than the pentagonal subunit used in the atomistic simulations, which hosted 10 molecules only.

As a control, the same simulations with external electric field were run with a DLPC membrane: for none of the values tested (0, 20 and 40 mV/nm) the capsule attached to the membrane, as already saw in standard MARTINI simulations.

Regarding diffusion, we retrieved the slowing down of lipids when the buckyball bilayer is bound to the bacterial model membrane. DLPG is more slowed down with

respect to DLPC in simulations with no or 40 mV/nm electric field, while for the 20 mV/nm it is slightly faster. Instead, when considering the DLPC membrane simulated in presence of a capsule (which does *not* bind to it), diffusion of lipids is increased with the field. This is the opposite to what observed for simulations of the DLPC membrane alone, showing that the buckyball bilayer has a long range interaction with the membrane and modifies its dynamics, even without binding to it.

As a final comment, it must be noticed that the variability on the diffusion coefficient might be higher than what suggested from the error derived from the fit. The MSD computed on MARTINI membranes has a very precise linear dependence from time, thanks to the length of the simulation and the extent of the system (the MSD is averaged over many time frames and many lipids, see Figure 3.42). This gives a small error on the fitted slope. However for the MARTINI simulation of the capsule on DLPC/DLPG, we run a second replica and observed a discrepancy of about 2 $\mu m^2/s$ between the diffusion coefficients, despite both have a very small error from the fit. As such, the results presented in Table 3.7 must be interpreted carefully. However, the differences between the DLPC and DLPG coefficients observed for the simulations after binding of the buckyball bilayer to the membrane are large enough to remain significant, confirming that the peptide has a stronger effect on negatively charged lipids.

3.6 Outlook

The investigation performed proved the power of the multiscale simulations approach to elucidate details of nanoscale systems which are otherwise inaccessible to experiments.

When investigating the assembly process of capzip molecules, the use of atomistic simulations gave insight into the role of each amino acid of the sequence in the pairing of many copies of the molecule. Coarse-grained simulations proved the stability of the hypothetical structure on the long time scale. Overall, we were able to conclude that the ability of capzip to form capsule lays not only in the scheme of opposite charges that it hosts along its arms, but also in the presence of many hydrophobic residues, which favour the assembly by hydrophobic effect. Thus, the proposed structure contains a

bilayer arrangement of the molecules, which is demonstrated to be more stable than a monolayer one, and is consistent with recent experimental findings showing multilayer capsules [Kepiro et al., 2019].

Coarse-grain simulations were also able to clarify the interaction of capzip with the membrane. Capzip has a propensity to bind to negatively charged membranes, which are a simplified model for the bacterial inner membrane, but not to zwitterionic lipids, which model a mammalian membrane. Moreover, on the bacterial membrane, the buckyball recruits the negatively charged lipids in its proximity, reducing their diffusion.

Atomistic simulations complemented these findings showing the details of such interaction: on a bacterial membrane, capzip interacts with the lipids forming many hydrogen bonds and inserting the Arginine side chains deep in the phosphate region of the lipids. Consequently, the lipids around the peptide are slowed down in their diffusion also on the time scale of atomistic simulations. When analogous simulations were run for the mammalian model membrane, less hydrogen bonds were observed, a smaller propensity for Arginine insertion, and a smaller reduction of the diffusion coefficient.

Finally, simulations with an esternal electric field applied to the system showed that the peptide bound to the bacterial membrane promotes poration at values of the field which would not cause electroporation on the membrane alone. Simulations at the atomistic resolution make clear that this process is initiated at the location of Arginine insertions, while a coarse-grained picture allows to see the deformation and partial opening of the buckyball passing through the membrane on longer time scales. Analogous experiments on the mammalian model membrane showed that its value for electroporation is lower than the one for the bacterial counterpart at the atomistic level, regardless the presence of capzip.

The above results integrated the ones deriving from experiments giving a more complete picture on the characteristics of capzip. However, it would be interesting to pursue the investigation further. Specifically, to extend the investigation of the assembly process from molecules dispersed in solution; to assess more extensively the capzip interaction with a mammalian model membrane; and finally to simulate more

complex systems, namely capzip in assembly with RNA to investigate its delivery ability. A brief outlook on these aspects will be given in Chapter 5.

3.7 Supplementary material

We include here additional material, plus Figures and Tables referenced in the chapter which can help the reader in interpreting the results.

3.7.1 Self-assembly simulations

Test simulations on the self-assembly properties of capzip were run with the standard MARTINI model. Indeed, the analysis presented in this chapter is a careful investigation of the determinants that can drive the assembly, but a vision of the process happening is still lacking. In Chapter 1 we already highlighted that simulating an assembly process requires a great amount of computational resources. For this reason, we focus on the MARTINI model with standard water, and we explore concentrations of the order of mM, which are much higher than the ones employed in experiments (around μM).

Results from 1.25 mM, 2.50 mM and 10 mM runs showed the formation of increasingly larger and interconnected clusters when raising the concentration. Respectively, the molecules aggregated into 7, 4 and 2 clusters within the first 6, 4 and 2.5 μs of the simulations (Figure 3.28). It is not excluded that longer simulation would result in all the clusters joined together.

The two properties we want to assess from these simulations are the contacts between the molecules, classified by amino acid type, and the SASA of each residue type, to be compared with the results from the capsule model. However, from the previous simulations we also know that the MARTINI parametrisation promotes a very tight assembly of the molecules once they get in proximity, and that this is partially reverted back when backmapping a structure to an atomistic description. Therefore, the results from self-assembly simulations must be interpreted with the caveat that the protein-protein interactions might be overestimated.

The backbone contacts (Figure 3.29) show between 3 and 4 contacts per residue,

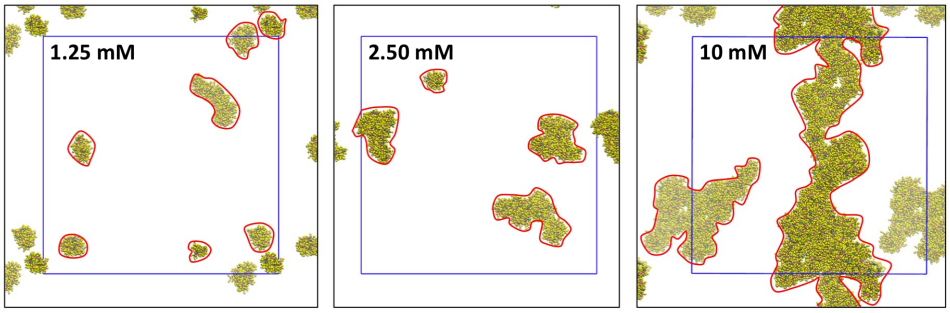


Figure 3.28: Final configurations obtained from 10 μs standard MARTINI simulations of the self-assembly of capzip molecules, from random initial configuration, at different concentrations. Boundaries of the simuation box in blue; different clusters are circled in red (the remaining ones are periodic copies of the ones highlighted).

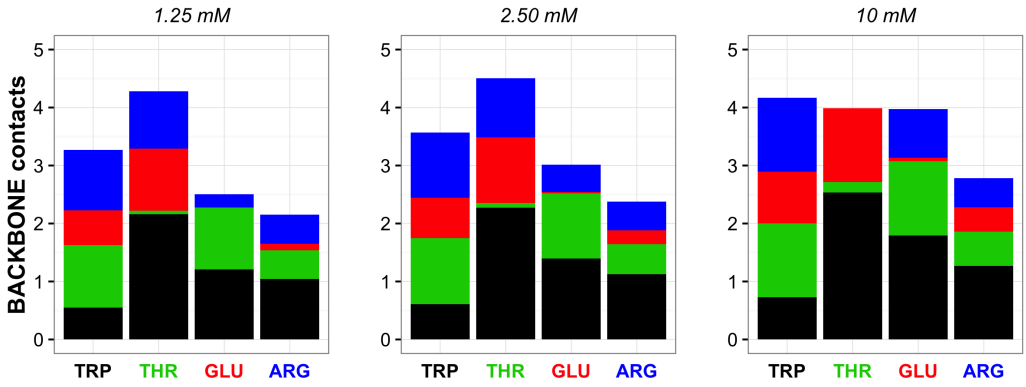


Figure 3.29: Number of contacts per residue type in each arm of capzip for a MARTINI self-assembly simulations at different initial concentrations: each bar shows the average number for the residue on the x-axis; its color is split by the identity of the partner residue (color coded as in the x-axis legend). For mixed contacts the residue on the x-axis contributes with its backbone. Only contacts existing more than 50% of the simulation time are considered.

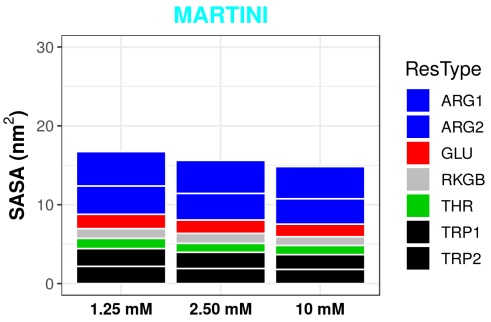


Figure 3.30: Solvent Accessible Surface Area (SASA) per molecule, divided by residue types, for self-assembly simulations at different concentrations.

which is about three times higher than the figure obtained for the ordered capsule structure (Figure 3.12), in line with a densely packed assembly. Furthermore, there is a slight increase with the concentration.

The pattern of contacts suggests that Threonine interact with most partners. Interestingly, the contacts are distributed in a way which resemble more the atomistic pattern found for the capsule rather than the MARTINI one. In general, this confirms the role of the hydrophobic residues in keeping the assembly compact, but also suggests aggregation of molecules when a high concentration is simulated.

Regarding SASA values (Figure 3.30), they confirm what found also in the capsule conformation, with Arginine being exposed and Tryptophan buried. The SASA is decreasing with the concentration, as for low concentration several clusters are formed and float in solution separately (up to the time scale simulated), while for higher concentrations the molecules can merge in fewer larger clusters, which have less exposed surface. Interestingly, the decrease in SASA and increase in contacts does not scale linearly with the concentration, as expected when the systems reach saturation and the clusters cannot favourably accommodate more molecules. However, as only three concentrations were simulated here, the data are not enough to extract a more precise dependence between these properties.

Again, when interpreting these results, one must remember that a backmapping to an atomistic resolution would produce more expanded structures, as it was observed for the capsule. However, also at the experimental level, high concentrations (of the order of mM as the ones simulated here) resulted in aggregation, i.e. the appearance of amorphous aggregates rather than geometrical capsules (data not shown). The MARTINI simulations are likely capturing this behaviour, and one would need a long simulation of a diluted solution to observe an ordered assembly. As it is almost impossible to simulate the micromolar regime used in experiments, the simulations of pre-assembled structures, as the ones presented previously in this chapter, are, so far, the only ones able to explore the relevant range of concentrations.

3.7.2 On the long range electrostatic cut off in lipid simulations

In a few paragraphs of this chapter we compare the simulations run here with the ones from the parametrisation work in the next chapter. As the reader will see, these are run on 512-lipid bilayers, with a simulation set up very similar to the one employed here.

However, the long range electrostatic are treated with a Reaction Field scheme and cut off radius of 1.4 nm. This choice was performed for consistency with previous parametrisation work.DLPC Instead, the simulations run in the present chapter are run either with the Reaction Field and a cut off of 1.2 nm (512-lipid bilayers) or with a Particle Mesh Ewald treatment, and the same cut off (740-lipid bilayers).

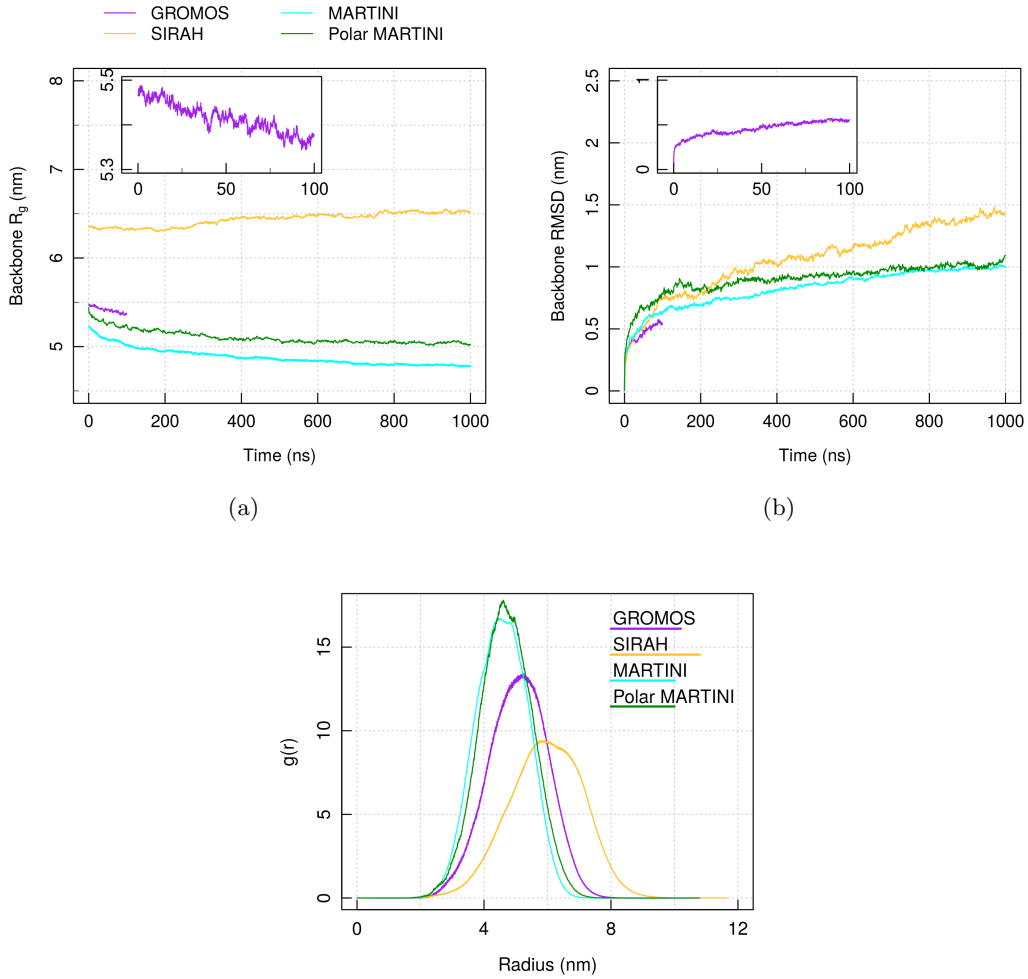
Although briefly mentioned in the body of the chapter, we want to reiterate that these discrepancies might affect the simulations, but in measure smaller than the peptide or the electric field, which effect we wanted to assess.

Indeed, extensive literature on the effects of long range electrostatic treatment suggests that electrostatics have an influence on the ApL up to 0.010-0.015 nm^2 only. This is confirmed for the change in the cutDLPC off treatment (single or double) [Silva et al., 2018; Reißer et al., 2017], and the switch from PME to RF [Poger and Mark [2012]; Table 1 in Chapter 4]. The only conditions which severely affects the ApL seems the complete neglect of the long range electrostatic, choosing a plain cut off scheme [Patra et al., 2003].

Regarding the different cut off values, we did not find previous work testing the difference. The effect of the cut off is to discriminate which regions are treated in the exact way (nearby the atom for which the electrostatic force acting on it needs to be computed) and which instead through the long range approximation of choice (far away ones). As these approximations seem consistent among them and robust, we foresee that a small shift in the cut off length does not have a high impact on the simulation outcome.

3.7.3 Buckyball simulations: results from Replica 2

In the following pages we include the results obtained from Replica 2 of the buckyball bilayer and monolayer simulations (analogous to Figures 3.9-3.14 and 3.16-3.18).



(a) (b)

(c)

Figure 3.31: (a) R_g and (b) RMSD computed on the Protein backbone. Results are displayed for simulations performed in GROMOS (100 ns), SIRAH, MARTINI and MARTINI with polar water (all 1 μs). Inset: zoom on the GROMOS values. (c) RDF of Protein masses around their center of mass, computed on both layers, displayed for the same simulation set up as in (a,b). For each label of the legend, the bar has length of the respective FWHM of the Gaussian function fitting the data (thickness estimate). Results for Replica 2 of each simulation set-up.

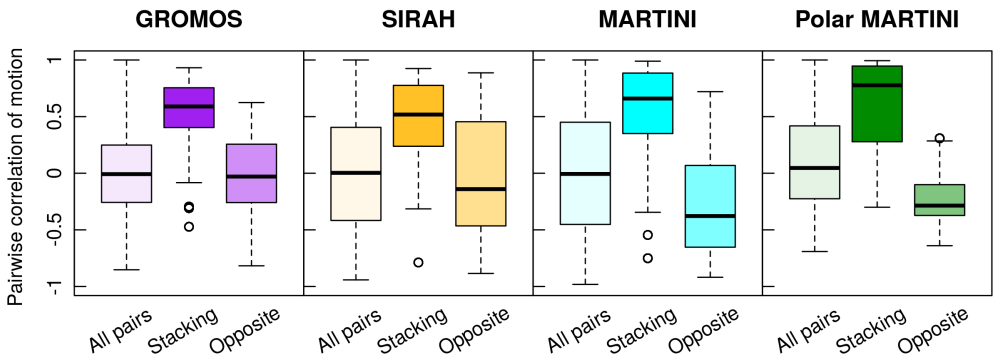
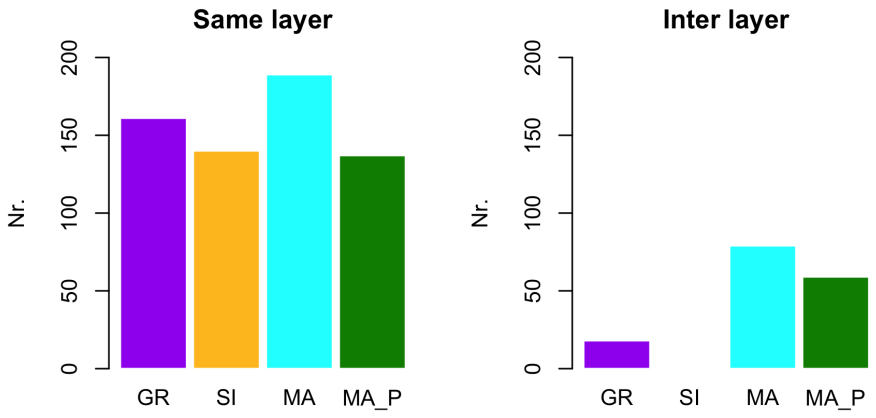


Figure 3.32: Distribution of the correlation of motion between different molecules in the buckyball simulations. Black band: median of the distribution; box: first and third quartiles; whiskers: maximum and minimum, outliers excluded (hollow dots). Results for Replica 2 of each simulation set-up.



DLPC

Figure 3.33: Number of paired arms within the same layer and between layers. Cut off distance between arms center of mass equal to 1.2 nm. Only contacts existing more than 90% of the simulation time are counted. Results for Replica 2 of each simulation set-up.

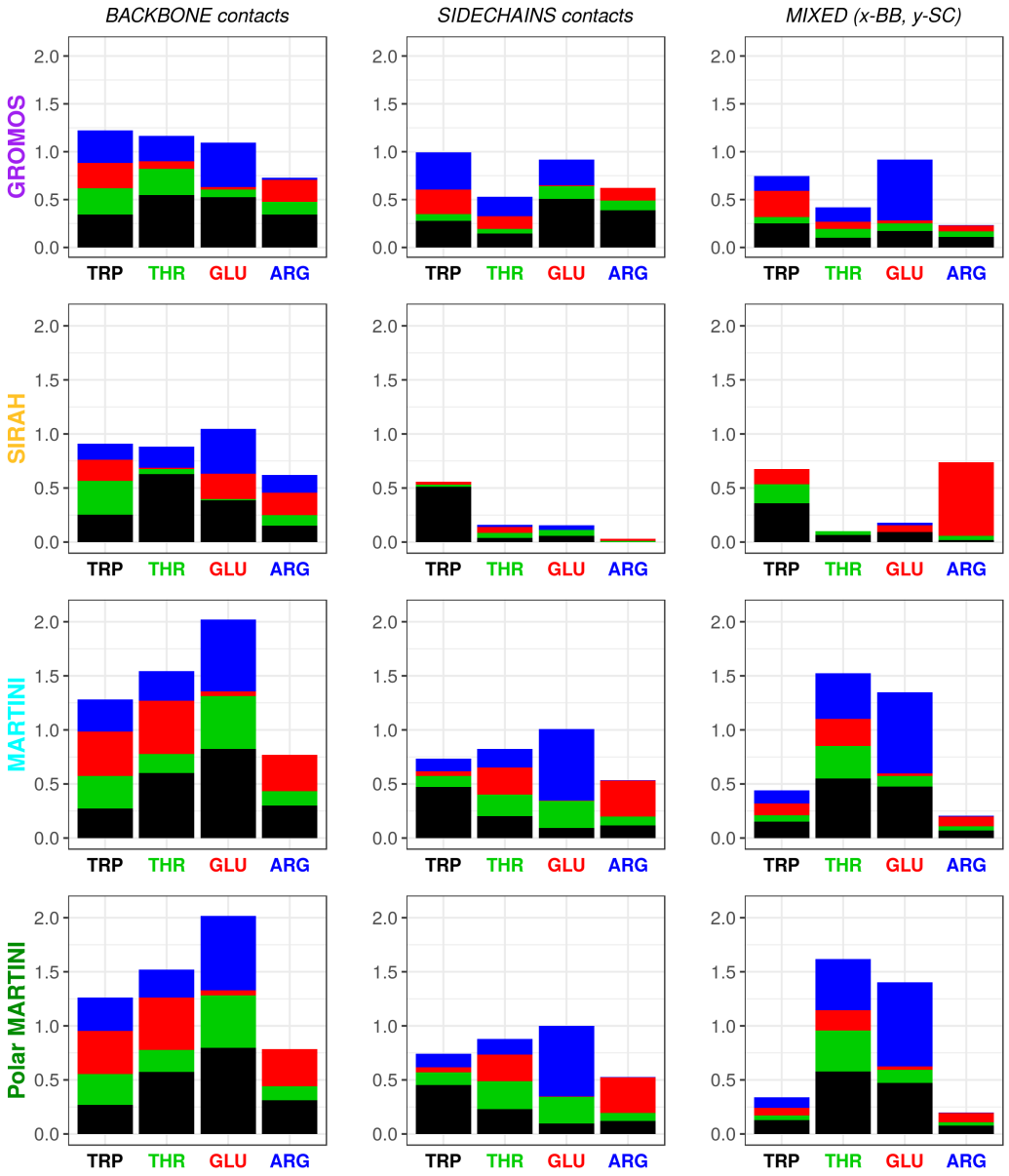


Figure 3.34: Number of contacts per residue type in each arm of capzip: each bar shows the average number for the residue on the x-axis; its color is split by the identity of the partner residue (color coded as in the x-axis legend). For mixed contacts the residue on the x-axis contributes with its backbone. The parametrisation is reported along the y-axis. Only contacts existing more than 50% of the simulation time are considered. Results for Replica 2 of each simulation set-up.

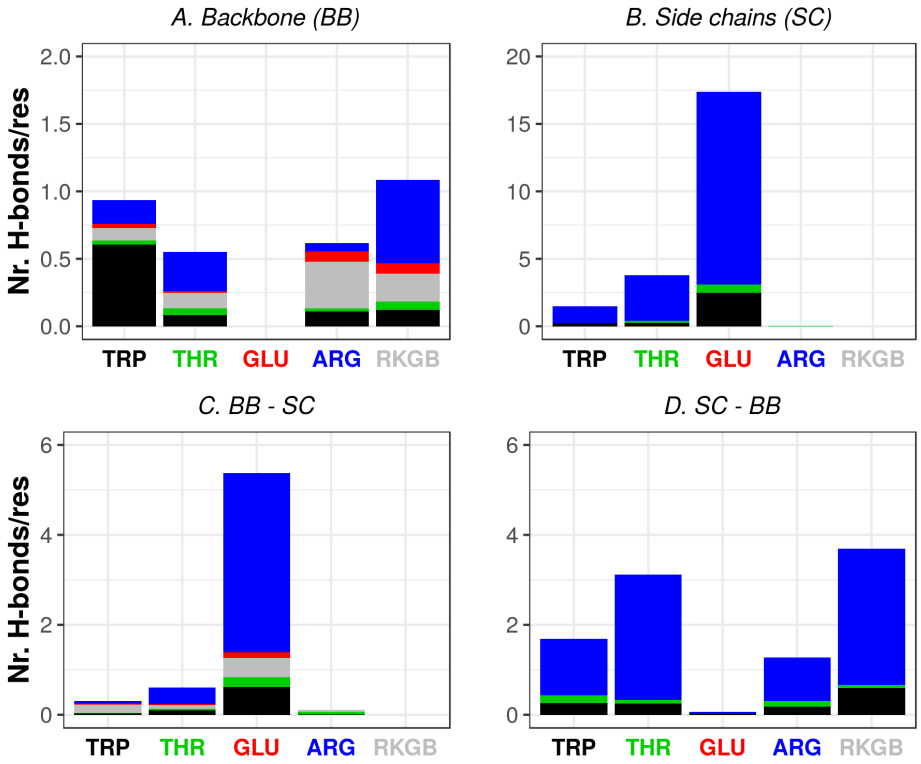
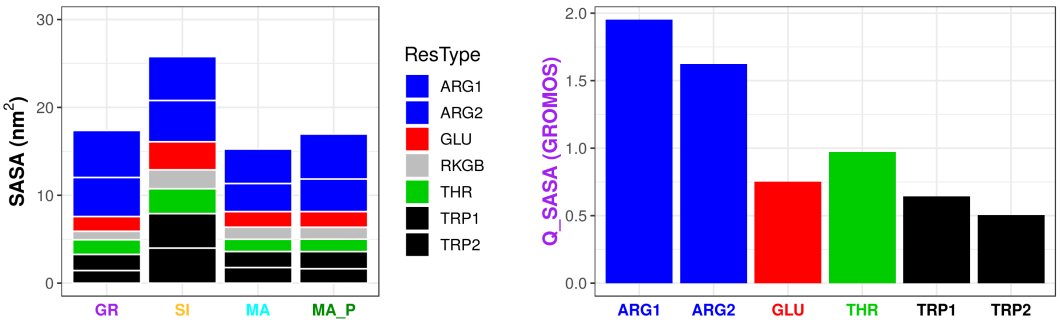
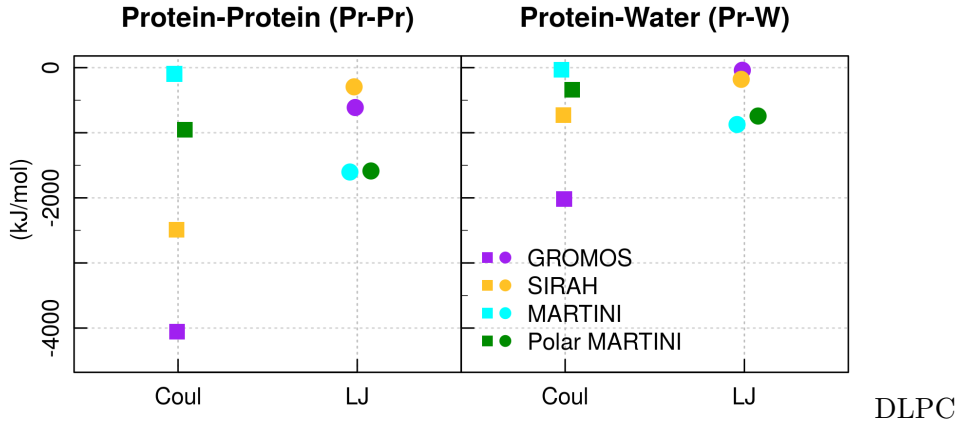


Figure 3.35: Average number of hydrogen bonds per residue occurring between amino acids, including the central scaffold RKGB, for a 100 ns atomistic simulation of the buckyball in solution. For each bar, the residue on the x-axis is the acceptor, and the bar is split by the identity of the donors. In the case of Backbone - Side chain and Side chain - backbone, the first mentioned correspond to the acceptor (and thus the residue on the x-axis). Results for Replica 2.

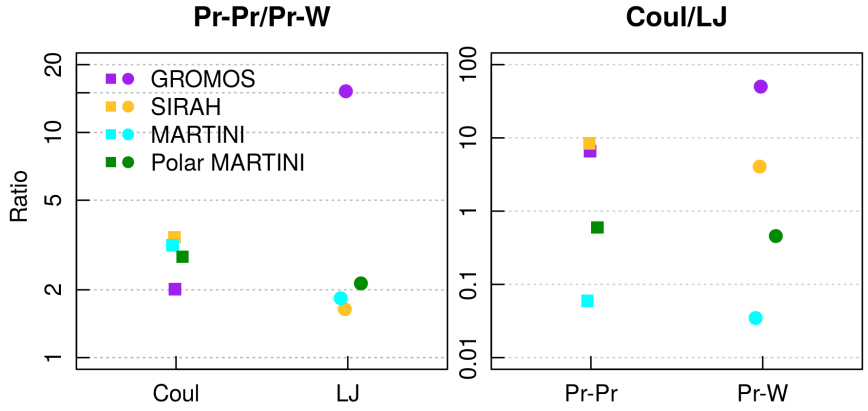


(a) (b)

Figure 3.36: (a) Solvent Accessible Surface Area (SASA) per molecule, divided by residue types. (b) Normalised SASA over the reference SASA computed for each amino acid type X as the value in a Gly-X-Gly tripeptide. Results for Replica 2 of each simulation set-up.



(a)



(b)

Figure 3.37: (a) Protein-Protein and Protein-Water non-bonded interactions, normalised per molecule. Values obtained as average on the second half of the trajectory of Replica 2 for each set-up. (b) Ratio between the Protein-Protein and Protein-Water interactions for each force field, for Coulomb and Lennard-Jones respectively; or between Coulomb and Lennard/Jones, for Protein-Protein an Protein-Water interactions separately (note the log scale on y). Values computed as for plot (a). Points are misaligned along x to facilitate the reading.

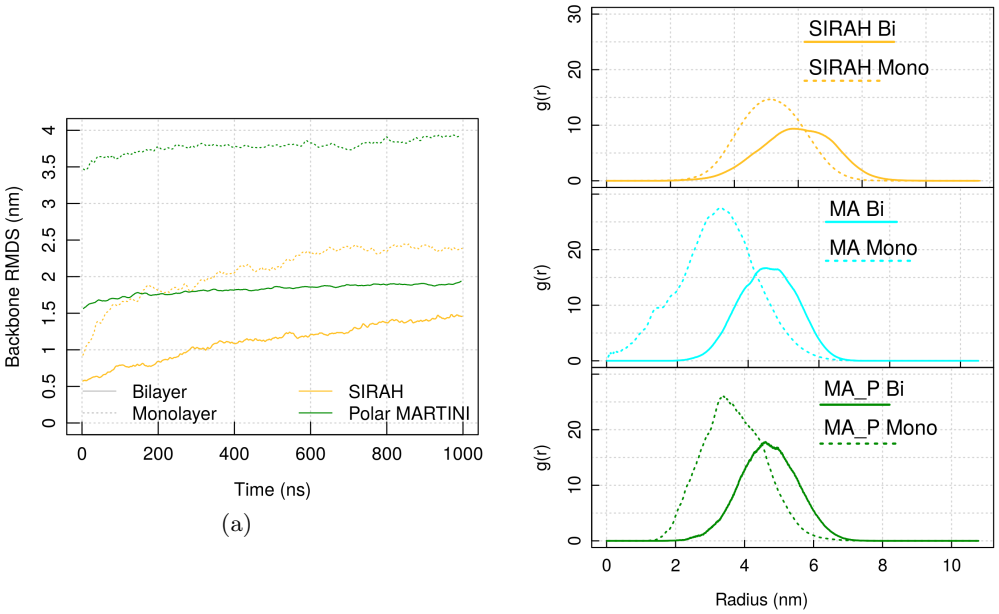


Figure 3.38: (a) RMSD of the monolayer and bilayer structures for SIRAH and Polar MARTINI force field, with respect to the initial geometrical configuration (external layer of Figure 3.1, E). (b) RDF of Protein masses around their center of mass. For each label of the legend, the bar has length of the respective RDF FWHM (thickness estimate). Results for Replica 2 of each simulation set-up.

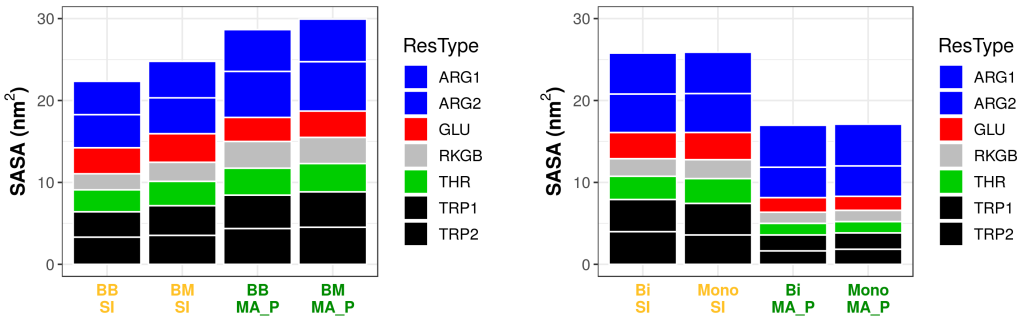


Figure 3.39: Solvent Accessible Surface Area (SASA) per molecule, divided by residue types for simulations of the bilayer and monolayer structure. (a) SASA computed from the initial configuration; (b) from the average over the production run. Results for Replica 2 of each simulation set-up.

3.7.4 Additional Figures and Table

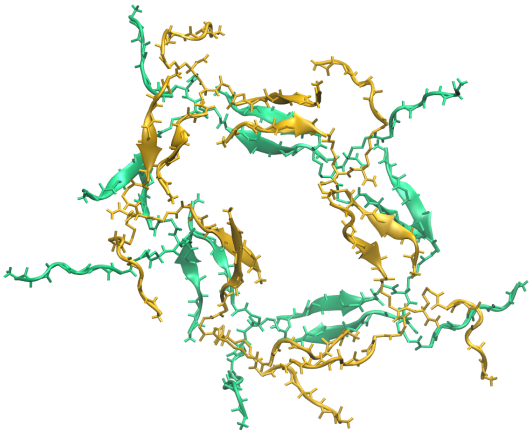


Figure 3.40: Final configuration from a 100 ns simulation of the pentagonal subunit in solution. Bonds and cartoon representation, coloured by name. [VMD software, Humphrey et al. [1996]]

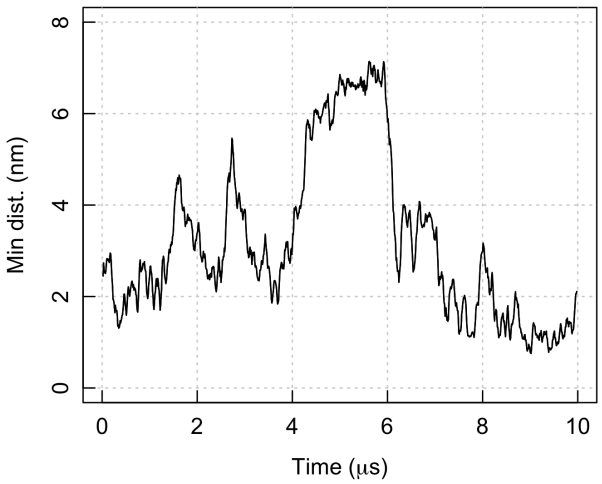


Figure 3.41: Minimal distance between the buckyball bilayer and a DLPC membrane (using the pair of closest atoms) during a MARTINI simulation.

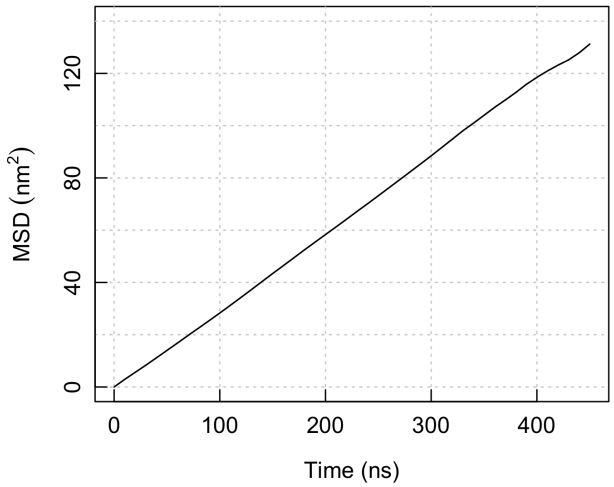


Figure 3.42: Typical lipid lateral MSD profile computed on Polar MARTINI simulations: MSD computed on the PO4 bead of DLPC for a 2880 lipids membrane, with an externally applied electric field of 20 mV/nm.

DLPCTable of simulations of pure membranes (atomistic)

Lipids	Box (nm)	E (mV/nm)	ES	Time (ns)	Rep.
United atom GROMOS (GR)					
512 (b)	12	0, 70, 120	RF	400	1^a
512 (b)	12	130	RF	600	3^a
512 (b)	12	140	RF	150^P, 154^P, 200	3^a
512 (b)	12	140	RF	200	1^b
740 (b)	14	0, 20	PME	400	1^b
748 (m)	14	0	PME	400	1^b
748 (m)	14	130	PME	20^P, 28^P, 39^P	3^b
Coarse-grained Polar MARTINI (MA_P)					
2880 (b)	30	0, 20, 40	PME	500	1
2888 (m)	30	0, 20, 40	PME	500	1
Coarse-grained MARTINI (MA)					
720 (b)	30	0	RF	1000	1
722 (m)	30	0	RF	1000	1

Table 3.8: Table of control atomistic simulations of membranes. All run at 150 mM concentration of NaCl. Lipids: number of lipids, (b) bacterial model (DLPC/DLPG 3:1) and (m) mammalian model (DLPC). Long-range electrostatic (ES): RF Reaction Field [Tironi et al., 1995], PME Particle Mesh Ewald [Essmann et al., 1995]. For the GROMOS force field (FF): ^{a} version 54a7 [Schmid et al., 2011], ^{b} 54a8 [Reif et al., 2012]; coarse-grained MARTINI [Marrink et al., 2007; Monticelli et al., 2008], coarse-grained MARTINI with polar water [Yesylevskyy et al., 2010]. Superscript P in the time length denotes observation of poration.

Chapter 4

Lipid parametrisation

Simulations of lipids can be a challenging task, especially when the membrane simulated has not been tested experimentally, and thus no comparison can be performed on standard properties such as the area per lipid. Moreover, the parameters describing each lipid species must be carefully chosen, to aim at an accurate reproduction of the natural properties of the membrane.

The work performed in Chapter 3 involved the simulation of a mixed membrane which, to our knowledge, has not been experimentally tested. Additionally, one of the lipids involved (DLPG) had not been parametrised before for the GROMOS force field. The procedure of DLPG parametrisation brought to our attention the discrepancy existing between the GROMOS parametrisation of protein and the one employed for lipids, as the two of them are obtained through different procedures: fit to hydration properties for proteins, and quantum mechanics computations for lipids.

The set of lipid parameters derived originally from Chiu et al. [1995] and updated multiple times up to the work of Poger et al. [2010], is still regarded as a standard for phosphocholines, and thus used to validate the most recent parametrisation of the force field (GROMOS 54a8 [Reif et al., 2012]). However, it does not take into account recent evolutions of the force field, aimed at reparametrising some constitutive moieties which happen to be also part of lipids. For this reason, we performed a reparametrisation of the lipid head group which includes these new descriptions, aiming at an improved consistency with the rest of the force field and monitoring whether they would also improve the agreement with the experimental quantities.

The resulting parametrisation was successful in reproducing key properties such as the area per lipid, and proposed a different interaction with a test peptide with respect to the one outlined by the previous parameter set. Specifically, the new parameters promote a weaker protein-membrane interaction allowing simulations to be less biased from the initial conditions chosen. This is of crucial importance for all simulations in general, and in particular when the sampling is reduced due to computational resources issues.

This chapter consists of the paper "Lipid Head Group Parameterization for GROMOS 54A8: A Consistent Approach with Protein Force Field Description", published in the *Journal of Chemical Theory and Computation*, together with its Supporting Information. The project was conceived by Prof. Franca Fraternali, Dr. Christian Margreitter and myself. I carried out all the work, under their guidance; specifically I implemented the new parameter sets, performed and analysed the simulations and selected the best performing parameters to be released. References for the paper and its supplementary material are included separately after each of them (with a numeric notation). All references feature in the bibliography of this thesis as well.

JCTC Journal of Chemical Theory and Computation
Article
Cite This: J. Chem. Theory Comput. 2019, 15, 5175−5193
pubs.acs.org/JCTC

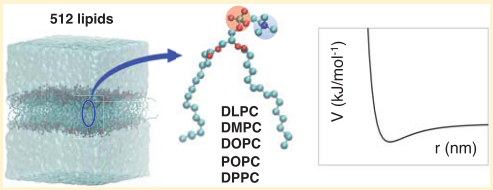
Lipid Head Group Parameterization for GROMOS 54A8: A Consistent Approach with Protein Force Field Description

Irene Marzuoli, Christian Margreitter, and Franca Fraternali*

Randall Centre for Cell and Molecular Biology, King's College London, London SE1 1UL, U.K.

Supporting Information

ABSTRACT: Membranes are a crucial component of both bacterial and mammalian cells, being involved in signaling, transport, and compartmentalization. This versatility requires a variety of lipid species to tailor the membrane's behavior as needed, increasing the complexity of the system. Molecular dynamics simulations have been successfully applied to study model membranes and their interactions with proteins, elucidating some crucial mechanisms at the atomistic detail and thus complementing experimental techniques. An accurate description of the functional interplay of the diverse membrane components crucially depends on the selected parameters that define the adopted force field. A coherent parameterization for lipids and proteins is therefore needed. In this work, we propose and validate new lipid head group parameters for the GROMOS 54A8 force field, making use of recently published parametrizations for key chemical moieties present in lipids. We make use additionally of a new canonical set of partial charges for lipids, chosen to be consistent with the parameterization of soluble molecules such as proteins. We test the derived parameters on five phosphocholine model bilayers, composed of lipid patches four times larger than the ones used in previous studies, and run 500 ns long simulations of each system. Reproduction of experimental data like area per lipid and deuterium order parameters is good and comparable with previous parameterizations, as well as the description of liquid crystal to gel-phase transition. On the other hand, the orientational behavior of the head groups is more realistic for this new parameter set, and this can be crucial in the description of interactions with other polar molecules. For that reason, we tested the interaction of the antimicrobial peptide lactoferricin with two model membranes showing that the new parameters lead to a weaker peptide−membrane binding and give a more realistic outcome in comparing binding to antimicrobial versus mammal membranes.



1. INTRODUCTION

Cellular membranes are key promoters and regulators of many biological processes due to their crucial role in segregating the external world from the organism. Small molecule transport, drug permeation, intracellular signaling, and antibody response are all regulated by the cell membrane or by membrane-related components.[1−8] To fully comprehend and ultimately influence the bespoke processes, it is paramount to understand membranes and their constituting lipids in atomistic detail. However, due to the complexity of those systems, researchers have resorted to the use of simplified model membranes, which can be synthesized and characterized in vitro. This enables the individual contributions of the components involved to be disentangled. Indeed, for the cellular membrane to be able to perform different functions, its composition is necessarily complex. Lipids are one of the main components and can be present in up to hundreds of different species.[9] In addition, many transmembrane proteins tessellate the cell surface, promoting signaling pathways and influencing the membrane's structural and mechanical properties.[10,11] Phospholipid bilayers and micelles have been investigated, in particular, as these lipids represent the main components of the eukaryotic and the inner bacterial membranes. Both have been modeled selecting specific phospholipids to emulate the appropriate surface charge or to reproduce the human cell membrane fluidity by introducing, for example, cholesterol.[12,13] As these simplified membranes retain the core characteristics of their different biological templates,[14] they can be used to test the membrane interaction with proteins, peptides, antimicrobial molecules, or drugs.

Experiments can provide global properties of membranes and, despite the great accuracy of techniques like NMR and X-ray scattering in measuring the average position of atoms in rigid structures, they face challenges when characterizing the biologically relevant fluid phase, as opposed to the gel one that emerges at lower temperatures.[15−18] Alongside experimental characterization, molecular dynamic (MD) simulations have played a central role in the investigation of the behavior of lipids, due to the atomistic spatial resolution they provide. Therefore, MD simulations complement our understanding of membranes' behavior and are also important for the study of lipid systems in combination with proteins, providing detailed insights into the mechanisms of their interactions. In the past,

Received: May 25, 2019
Published: August 21, 2019

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

Journal of Chemical Theory and Computation Article

MD simulations have been successfully employed to reproduce typical phenomena in membranes, such as lipids' flip-flop,[19,20] vesicle formation,[21,22] aggregation into bilayers,[23−25] and stress-induced[26−29] and peptide-induced pore formations.[30−32] Moreover, the implementation of more realistic models of bacterial membranes, by including a more diverse set of components into the simulated systems, has been pursued[33,34] to test specific interactions with antimicrobial peptides and understand their selectivity.[35,36]

The reliability of such simulations depends on the accurate parameterizations of lipids and proteins, which need to be validated against experimental data. Moreover, the two descriptions must be consistently integrated into the force fields used, i.e., be derived with the same parameterization procedure. Different approaches to the problem are possible, which resulted in the development of multiple force fields suitable for simulations of biomolecules: for example, the CHARMM[37−39] and AMBER[40] force fields are parameterized from quantum mechanics calculations, while GROMOS96[41] is calibrated to match global properties like the hydration free energy of chemical moieties. All of them have been constantly updated to meet the new experimental values available and more faithfully reproduce the different species involved.

However, it is a very difficult task to parameterize the constituents of a complex system so that all parameters are consistent with the rest of the force field and reproduce both the single-molecule observables and the collective behavior. In the present work, we consider the parameterization of phospholipids in the context of the GROMOS96 force field,[41] addressing some of the inconsistencies in the lipid head group parameters commonly used so far, particularly in consideration that these contribute to the description of recognition processes at the interface.

In the past, lipid simulations using the GROMOS96 force field suffered from difficulties involved in transferring the pre-existing parameters, calibrated mainly for peptides in an aqueous environment, to the amphiphilic environment of the lipid assembly. This resulted in the failure to reproduce the membranes' behavior properly[42−44] and therefore a series of modifications were adopted, particularly in the choice of lipid-specific Lennard-Jones interactions[44−46] and partial charges.[42]

In the light of recent reparameterizations of a set of choline moieties[47] and of phosphate-containing species,[48] we undertake the task of updating the parameters used for lipids, in particular, phosphocholines, as they contain both these chemical moieties. Within this work, we show that it is possible to integrate the recently computed partial charges within simulations while maintaining good agreement with the available experimental data. We also test the transferability of the new phosphate charges onto lipids without a choline head group, namely, phosphoethanolamine (POPE) and phosphatidylglycerol (POPG).

Most importantly, the new description of phosphocholine is consistent with the GROMOS96 parameterization philosophy, based on the decomposition of large molecules into smaller compounds and subsequently fitting their parameters to experimental hydration free energies. Together with adjustments to specific van der Waals potentials, we believe that the parameters presented here will contribute to improving the accuracy of the description of membrane−solvent and membrane−protein interactions. To this aim, we compared the available parameters with the one proposed in this work, simulating the interaction of an antimicrobial peptide with two model membranes, highlighting the differences in the mechanisms observed, and comparing them with the available experimental evidence.

2. METHODS

2.1. Background to Lipid Force Fields. The most recent iteration of the lipids' parameters commonly used in simulations with the GROMOS force field is the one by Poger and Mark.[44] They employed partial charges derived quantum-mechanically by Chiu et al.,[42] combined with a modified repulsion between the choline methyl groups and the OM oxygen atoms in the phosphate with respect to the standard choline−OM one.

The original set of Chiu charges[42] was derived from ab initio Hartree−Fock self-consistent field calculations[52] and Mulliken population analysis.[53] Slight modifications were applied to make each individual charge group sum up to an integer value, following the GROMOS96 philosophy. Despite the resulting charges that differ substantially from the ones used for the same chemical groups in different chemical contexts, the GROMOS community employed this set as it gave results in closer agreement with the available experimental data.

The refinement of van der Waals parameters for aliphatic alkanes, together with the bond, bond angle, and torsional parameters for the ester groups,[49,50] prompted the reparameterization of the lipids' head group description: in particular, Chandrasekhar et al.[45] recomputed the head group torsional parameters from ab initio quantum-mechanical torsional profiles of each of the fragments composing the head group (Figure 1).

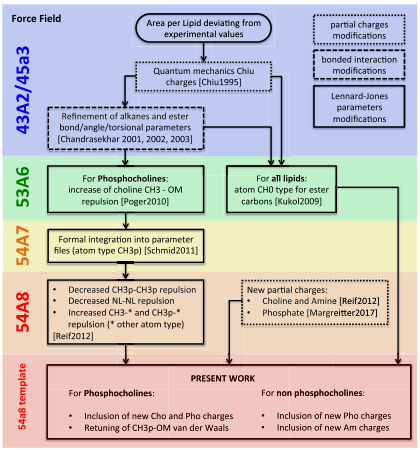


Figure 1. Evolution of the lipid parameters in the GROMOS force field. References—Chiu1995: ref 42; Chandrasekhar2001, 2002, 2003: refs 45, 49, 50; Poger2010: ref 44; Kukol2009: ref 46; Schmid2011: ref 51; Reif2012: ref 47; Margreitter2017: ref 48.

The modifications above were included in the 53A6 version of the GROMOS force field. However, an additional change was necessary to match the experimental results. Therefore, Poger and Mark[44] introduced a change in the CH3 choline and OM repulsion. The C12 parameter (related to the Pauli repulsion) between the newly introduced atom-type CH3p (to

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

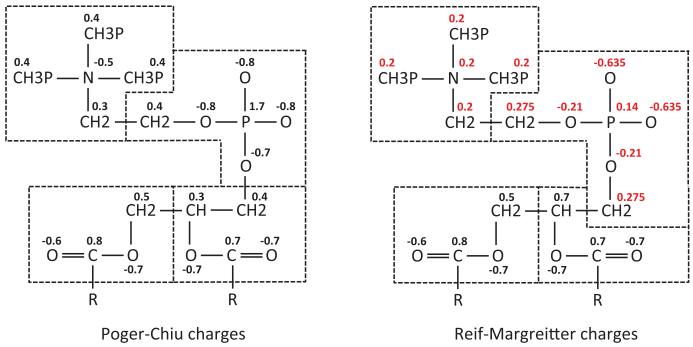


Figure 2. Partial charges for the phosphocholine head groups and the glycerol and ester moieties in the Chiu[42] scheme (left) and the one tested in the current work (right). Red font denotes values that have been changed between the two. Atoms belonging to the same charge groups are enclosed by the same dashed polygon.

Table 1. Table of Simulations for Phosphocholine Bilayers[a]

simulations of phosphocholine lipids				
sim	charges[b]	FF	CH3p−OM C12[c] (kJ mol^{-1} nm^{12})	CH3p−CH3p C12[d] (kJ mol^{-1} nm^{12})
1	Chiu	54A7	1.58 × 10^{-5}	2.66 × 10^{-5}
2	Chiu	54A8	6.93 × 10^{-6}	6.48 × 10^{-5}
3	RM	54A8_v1	1.10 × 10^{-5}	6.48 × 10^{-5}
4	RM	54A8_v2	1.58 × 10^{-5}	6.48 × 10^{-5}
5	RM	54A8_v3	4.50 × 10^{-5}	6.48 × 10^{-5}

[a]All are run for 500 ns and systems consisting of 512 lipid molecules (256 per layer), using a particle mesh Ewald (PME) long-range electrostatic scheme. [b]Charge set: Chiu from Ref 42, Reif−Margreitter (RM) as illustrated in the present work. [c]As a reference, the standard C12 parameter in 54A7/54A8 for CH3−OM is 4.44 × 10^{-6} kJ mol^{-1} nm^{12}. [d]The CH3−CH3 C12 parameters are 2.66 × 10^{-5} for each parameter set.

represent the united methyl atoms in the choline head group of lipids) and the oxygen-type OM, present in the phosphate group, was increased by a factor of 3.5. This modification was optimized and tested against experimental values, increases the spacing between individual lipids, and thus leads to the appropriate area per lipid (ApL).[51,54] The new atom-type CH3p has all of the characteristics of CH3, except for the bespoke parameter, i.e., the Lennard-Jones interactions involving OM. These Poger−Chiu parameters have been successful in reproducing membrane behavior and were used in many MD applications.[55−57]

Later, Reif et al.[47] enhanced the methyl−methyl repulsion for both CH3 and CH3p in the 54A8 parameter set, which allowed for a decrease in the large repulsion value between CH3p and OM previously introduced[44] while still reproducing experimental values. The 54A8 parameter set contains two additional, nonlipid-specific, modifications important for this work: the choline Lennard-Jones parameters and partial charges, and the phosphate partial charges. The C12 Lennard-Jones repulsion term for the NL nitrogen atom type (present in the choline moiety) was increased to successfully prevent oversolvation.[47,54] To the same end, the +1 e total charge was evenly distributed over all five atoms, which resulted in a better approximation of the experimentally obtained hydration free energy in comparison to the 54A7 parameter set. Similarly, Margreitter et al.[48] calibrated the partial charges of four phosphate species and enhanced the reproduction of experimental data. The relevant phosphate-containing species for this work is dimethyl-phosphate, a compound not directly present in force field versions prior to 54A8.

Another approach to lipid parameterization was proposed by Kukol,[46] namely, the use of the already available CH0 atom type for the ester carbons in place of the standard C atom type, in conjunction with the Chiu charges. This atom type, designed to describe a bare sp^3 carbon bound to four heavy atoms, has a repulsion energy term 10−40 times larger than a bare carbon bound to other atom types, enforcing a greater spacing between lipid molecules and thereby increasing the ApL. As this modification is also applicable in the absence of a choline head group and does not require the introduction of another atom type, this method can be used to parameterize POPE and POPG.

2.2. Parameterization Strategy. In an effort to enhance the consistency of the force field, we integrated the new partial charges for the choline and phosphate moieties [Reif−Margreitter (RM) charge set] into the lipid building blocks of GROMOS 54A8 so that the entire phosphocholine head group now follows the common GROMOS-like modeling approach (Figure 2). Only the partial charges of the ester groups remain as described in the Chiu set, a deviation from the canonical parameterization strategy necessary to match the experimental area per lipid values: Chandrasekhar et al. showed in ref 59 that the replacement of the ester charges with the standard ones for the ester moiety (parameterized to reproduce the experimental free energies of hydration of a series of alkane esters[60]) resulted in a much smaller area per lipid, not compatible with the experimental values.

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

The introduction of the new head group charges required a refinement of the CH3p−OM Lennard-Jones repulsion, as the 54A8 value was set considering the original Chiu charges. Ideally, one would always try to keep the force field terms as much transferable as possible. Nevertheless, the complexity and anisotropic nature of some biological environments can be difficult to parametrize with single chemical groups, as the same chemical group can behave differently according to the context it is inserted in. Lipid systems are one of such examples, and to maintain the correct physical behavior of the system, we used specific C12 parameters for the CH3p−OM repulsion in phosphocholine lipid atoms. This allows for a more balanced description of the physicochemical properties of the lipid bilayer and a better match with the available experimental observables.

Aiming at this, and to disentangle the effect of charge parameterization versus the CH3p−OM repulsion, we tested three different values of such Lennard-Jones parameter with the new charges while control simulations were run using the Chiu partial charges and the GROMOS 54A7 or 54A8 parameter set for each lipid (Table 1). The phosphocholines tested are 1,2-lauroyl-*sn*-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC), 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC), and 2-oleoyl-1-palmitoyl-*sn*-glycero-3-phosphocholine (POPC), which have different tail lengths and numbers of unsaturated bonds as in previous works[44,61] (Table 2).

Table 2. Details of the Systems Simulated: Lipid Name, Tail Composition, Initial ApL (and Reference from Which the Initial Coordinates Are Taken), Simulation Temperature, and Gel−Liquid Phase Experimental Transition Temperature

lipid bilayer systems				
lipid[a]	tails[b]	ApL_0 (nm^2)	T_{MD} (K)	T_C (K)
DLPC	12:0/12:0	0.632[44]	303	276.4[66−69]
DMPC	14:0/14:0	0.616[61]	303	296.9[66−69]
DOPC	18:1c9/18:1c9	0.649[61]	303	255.7[66−69]
POPC	16:0/18:1c9	0.638[61]	303	270.5[66−69]
DPPC	16:0/16:0	0.631[61]	323	314.2[66−69]
POPE	16:0/18:1c9	0.568[33]	313	299.3[70]
POPG	16:0/18:1c9	0.602[33]	303	268.1[71]

[a]DLPC: 1,2-lauroyl-*sn*-glycero-3-phosphocholine, DMPC: 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine, DOPC: 1,2-dioleoyl-*sn*-glycero-3-phosphocholine, POPC: 2-oleoyl-1-palmitoyl-*sn*-glycero-3-phosphocholine, DPPC: 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine, POPE: 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine, POPG: 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1′-*rac*-glycerol). [b]Example: 16:0/18:1c9 indicates that tail 1 has 16 carbons with no unsaturated bonds and tail 2 has 18 carbons with one unsaturated bond between carbons 9 and 10—ester carbon counts as number 1.

To prove the transferability of the new phosphate charges to other lipid species, which do not contain a choline head group (and thus an enhanced repulsion, which has an impact on the ApL), test simulations of a phosphoethanolamine (POPE) and a phosphoglycerol (POPG, Table 2) bilayer have been performed. These lipids have amine and glycerol head groups, respectively. The parameterization of both takes advantage of the Kukol approach[46] employing a CH0 atom for the ester moieties to enhance the repulsion between lipids. For POPE and POPG, simulations were run with the standard parameters from ref 33 (denoted as Piggot−Chiu in the present work) or with the updated phosphate partial charges (Supporting Information (SI) Table 2).

The evolution of simulation techniques seen in the recent years suggested two other changes in the simulation setup: first, the original set of parameters was designed to be used with a twin-range cutoff scheme and a reaction field long-range electrostatic contribution,[62] but the twin-range cutoff is no longer supported in the latest versions of the GROMACS software used for the present work.[63] Additionally, the PME algorithm[64] for long-range electrostatic treatment is currently the predominant method used for protein dynamics. In the context of unifying the two fields of protein and lipid simulations, we therefore opted for a PME long-range treatment, running a control simulation (on the DPPC bilayer) with a reaction field scheme to assess the impact of such a change (SI Table 1).

The other change we adopted in comparison to the earlier work was a larger system size. Due to computational limitations, the original parameterization was performed on a 128-lipid bilayer,[44] but recent advances allow for larger systems to be simulated and we therefore used membranes four times as large (512 lipids). This larger size allows to track larger undulations of the membrane, as the effect of periodic boundary conditions (PBCs) is less restrictive. Again, a control simulation on a 128 DPPC membrane has been run to test the relevance and the effect of this change (SI Table 1).

Finally, the improvements reached with the adoption of the new parameters are monitored through the comparison with experimental values, but it is useful to have benchmarks derived from other simulation experiments. For that purpose, we compare some key properties with the values obtained by the all-atom CHARMM36 force field.[37,65] Despite a thorough comparison is beyond the present work, it is relevant to observe whether the changes introduced by the new parameters are going in the direction of the outcomes proposed by other descriptions.

2.3. Simulation Systems. Seven pure lipid bilayers have been simulated, five of which contain phosphocholines, one phosphoethanolamine (POPE), and one phosphoglycerol (POPG), as described in Table 2. Every bilayer is formed by 512 lipids (256 per leaflet), generated by replicating an equilibrated 128-lipid system from the literature two times in the x and y directions (see Table 2).

Water molecules were added to reach a minimum distance of 7.5 nm between periodic copies of the membrane along the z-direction, with a ratio of 85−120 H_2O per lipid. This distance is larger than the one used in the previous parameterization publications because we observed an enhanced undulatory behavior for larger membranes and therefore a higher distance is necessary to avoid interactions between periodic replicas in the z-direction.

2.4. Simulation Parameters. All simulations were run using the GROMACS software version 2016.3,[63,72,73] under periodic boundary conditions in a rectangular box. The temperature was maintained by coupling the membrane and the solvent independently to an external bath using the Berendsen thermostat[74] with a coupling time τ_T of 0.1 ps, at the reference temperatures indicated in Table 2, which are above the gel−liquid phase transition temperature for each lipid. The pressure was kept at 1 bar with a semi-isotropic coupling using a Berendsen barostat,[74] applying isothermal

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

compressibility of 4.5 × 10^{-5} bar^{-1} and a coupling constant τ_P of 0.5 ps. Covalent bond lengths of the lipids were constrained using the LINCS algorithm.[75] The geometry of the simple point charge water molecules was constrained using SETTLE.[76] A 2 fs time step was used, with a Verlet integration scheme. The PME[64] long-range treatment was applied to the electrostatic interactions beyond a 1.4 nm cutoff, and the reaction field scheme[62] control simulation was run with the same cutoff radius. A plain cutoff was used for van der Waals interactions, with a cutoff radius of 1.4 nm.

Each system was initially energy-minimized and then simulated at 50 K for 10 ps. Subsequently, the temperature was increased gradually over 500 ps until the final simulation temperature. The system was then simulated for 500 ns. The equilibration of the systems was monitored by examining the time evolution of the potential energy and the area per lipid: 200 ns is found to be sufficient to reach equilibration for all of the bilayers (SI Figure 2) so that the analysis has been performed over the last 300 ns of the production run, with frames stored every 100 ps. An overview of the simulations performed is given in Table 1 and SI Tables 1 and 2.

2.5. Analysis. To calibrate the lipid parameters, we used the observables listed below, as common practice in standard parameterization procedures.[61,77]

2.5.1. Area per Lipid. For systems where the membrane is aligned to the *xy* plane, the area per lipid (ApL) can be computed from the product of the lateral dimensions of the simulation box divided by the number of lipids in one leaflet. As shown in SI Figure 2 for DPPC, after 100 ns of simulation, the ApL oscillates around a value with fluctuations of the same magnitude, indicating equilibration. To allow further time for local rearrangements, we restrict our analyses to the last 300 ns of the simulations.

The equilibration protocol was verified on the DPPC bilayer, repeating the computation of the ApL on two nonoverlapping time windows, specifically between 200 and 350 ns and between 350 and 500 ns. For all of the parameter sets, the two windows gave compatible values of the ApL, confirming the convergence of the simulations (SI Figure 3).

The above procedure is valid if the membrane is flat or has minor undulations only. To test this and verify that deviations from planarity are not influencing the results, the ApL was recomputed for DPPC taking into account membrane undulations according to the procedure outlined in ref 78. The differences with the values computed from the simulation box dimensions were between 0.20 and 0.46%, which is lower than the error derived from the standard deviation across the simulation for any of the area per lipid computed.

As such, our computations are of value in rating the results against experimental outcomes and/or to compare parameter sets, as a local measure would not significantly improve the comparison.

2.5.2. Isothermal Area Compressibility Module. Following the protocol in ref 61, we computed the isothermal area compressibility module (K_A) from the fluctuations of the ApL values according to

$$K_A = \frac{k_B\langle T\rangle\langle \mathrm{ApL}\rangle}{n_L\,\sigma_{\mathrm{ApL}}^2} \tag{1}$$

where k_B is the Boltzmann constant, $\langle T\rangle$ and $\langle \mathrm{ApL}\rangle$ are the ensemble averages of the temperature and the area per lipid, respectively, n_L is the number of lipids in one leaflet, and σ_{ApL}^2 is the variance of ApL.

2.5.3. Bilayer Thickness. From the electron density profiles, the bilayer thickness can be evaluated in several ways and compared to the values from X-ray scattering experiments: the hydrophobic thickness (D_{HH}) is measured as the distance between the phosphorus peaks in the two layers, as these atoms have the highest electron density, while the Luzzati thickness (D_B)[61] is defined as

$$D_B = b_z - \int_{-b_z/2}^{+b_z/2} \rho_W(z)\,\mathrm{d}z \tag{2}$$

where b_z is the *z*-dimension of the simulation box and $\rho_W(z)$ is the volume fraction of water (vs other components) along *z* and normalized to 1 in the bulk water region

$$\rho_W(z) = \frac{n_W(z)V_W}{\mathrm{d}V} \tag{3}$$

where $n_W(z)$ is the time-averaged number of water molecules in a bin of width d*z*, V_W is the specific volume of the water model used (taken from ref 79), and d*V* is the time-averaged volume of a slice.

2.5.4. Dipole Potential. The dipole potential along the *z*-direction (perpendicular to the membrane plane) can be computed from the charge density along *z* ($\rho(z)$) via a double integration[82]

$$\psi(z) = -\frac{1}{\epsilon_0}\int_{z_0}^{z}\int_{z_0}^{z'} \rho(z'')\,\mathrm{d}z''\,\mathrm{d}z' \tag{4}$$

Several choices are possible for the two integration constants,[83] and for the present work, they are selected to set the dipole potential to zero in the middle of the bulk water region, at both sides of the membrane.

2.5.5. Deuterium Order Parameter of Lipid Chains. The deuterium order parameters S_{CD} of the acyl chains for each lipid bilayer were calculated and compared between the different sets studied. S_{CD} evaluates the average order of the lipid tails by measuring the orientation with respect to the bilayer normal of a carbon−hydrogen bond in a given position along the chain for each lipid in the bilayer. Their spread is evaluated according to the ensemble average

$$S_{CD} = \frac{1}{2}\langle 3\cos^2\theta - 1\rangle \tag{5}$$

As the GROMOS force field employs a united-atom representation, the tetrahedral positions of the hydrogens are constructed based on the neighboring carbons' positions.[58,80,81]

2.5.6. Hydration of Head Groups. To estimate and compare the hydration of lipid molecules, we computed the distribution of the distances between the oxygen of water and the nearest lipid atom. For each individual chemical group, the distance between the water oxygen and the nearest atom within that group was calculated. A quantitative measure for hydration was obtained by integrating the distribution up to the first peak or second one (for phosphate and glycerol).

2.5.7. Orientation of Head Groups. We computed the orientation of the lipids' head groups as the distribution of the angle between the P−N vector (joining the phosphorus atom and the choline nitrogen) and the outward normal to the membrane. The orientation of the *sn*-1 and -2 carbonyl dipoles with respect to the bilayer normal has also been calculated.

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

Journal of Chemical Theory and Computation Article

Table 3. Average Area per Lipid (in nm^2) over the Last 300 ns of Simulations for Phosphocholine Bilayers[a]

		ApL (nm^2)				
ID	charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	0.608(4)	0.591(4)	0.600(5)	0.604(4)	0.616(4)
2	Chiu/54A8	0.626(5)	0.612(4)	0.623(4)	0.623(4)	0.635(5)
3	RM/54A8_v1	0.631(4)	0.616(5)	0.625(6)	0.629(5)	0.638(5)
4	RM/54A8_v2	0.652(5)	0.643(6)	0.649(5)	0.650(5)	0.657(5)
5	RM/54A8_v3	0.690(6)	0.684(6)	0.690(6)	0.687(6)	0.693(5)
0	RM/54A7					0.603(5)
RF	Chiu/54A7					0.603(4)
small	Chiu/54A7					0.594(11)
experimental[b]		0.608−0.632	0.589−0.660	0.674−0.725	0.643−0.683	0.570−0.717
CHARMM36[37]		0.644(4)	0.608(2)	0.690(3)	0.647(2)	0.629(3)

[a]The number in parentheses is the standard deviation of the last digit. All simulations are run at 303 K, except for DPPC (run at 323 K). Analogous values for POPE and POPG are reported in SI Tables 6 and 8. [b]We report the maximum and minimum values of a review of experimental results given in Table 1 of ref 77. Only values referring to the temperature simulated are considered.

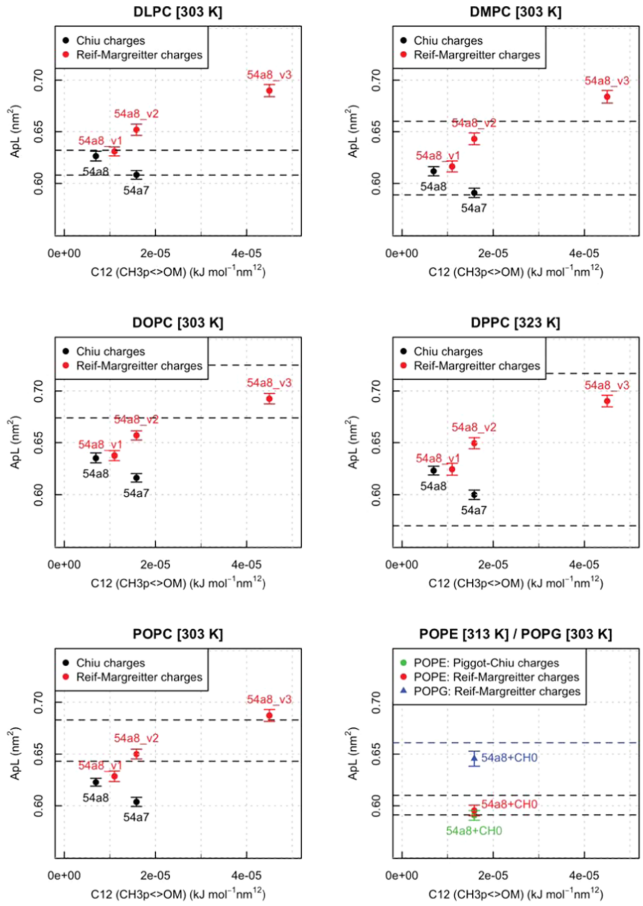


Figure 3. Area per lipid obtained for the five sets of parameters and seven lipid species. Error bars are the standard deviation over the 300 ns analyzed. Dashed lines indicate the range of experimental values from Table 1 in refs 77 and 33. For the plot reporting POPE and POPG values, the black dashed lines refer to POPE and the blue one to POPG.

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

2.5.8. Lateral Diffusion. For each simulation, we extracted the trajectory of the phosphorus atom of every lipid in the top and bottom leaflets separately, removing the collective motion of the leaflet. These trajectories were used to compute the mean-square displacement (MSD) for each lipid as a function of time, discarding the first 200 ns of production. This figure was averaged over all of the lipids in the leaflet and, for a given interval of time, on all of the possible time windows of that length-fitting within the simulation time analyzed. The diffusion coefficient D was obtained from a linear fit of the average MSD profile, following the Einstein equation[84] in two dimensions

$$\langle (x - x_0)^2 \rangle = 4Dt \tag{6}$$

The fit was performed discarding the first 50 ns of the profile, where the behavior is not linear, and the last 100 ns, where the poorer statistics leads to more noisy data. Coefficients obtained for the two leaflets were averaged to give the value reported.

2.5.9. Tilt Modulus. We computed the tilt modulus following the theoretical framework explained in ref 85. According to this, the angle θ a lipid forms with the local normal to the membrane follows the distribution

$$P(\theta) = C \sin(\theta) \exp\left(-\frac{\kappa_t^1 \theta^2}{2k_B T}\right) \tag{7}$$

where C is a normalization constant, k_B is the Boltzmann constant, T is the temperature, and κ_t^1 is the tilt modulus. This can be extracted from a fit of the distribution or, for computational reasons, from a fit of $\ln(P(\theta)(\sin\theta)^{-1})$. The direction in which a lipid points is taken as the vector joining the center of mass of the terminal atoms of the tails and the center of mass of selected atoms in the head group. Specifically, the last three carbons of each tail are taken as the reference for the first group, and the phosphorus and the carbon from which the two tails divert for the second. The computation was performed using a dedicated python module[85] available on the openStructure platform.[86]

2.6. Phase Transition. The set of new parameters performing best according to the previous observables was tested for sensitivity to temperature variations. A DPPC bilayer was chosen as the reference system and simulated at two additional temperatures: 303 and 333 K (SI Table 1), the first of which is below the experimentally determined liquid to gel-phase transition temperature.[66−69] As DPPC has also been used to perform the other control simulations, we opted for this model membrane for consistency reasons.

Besides the standard analysis described before, the local area per lipid was computed using a Dirichlet tessellation[87] of the lipid tail positions projected onto the horizontal plane parallel to the membrane (one leaflet at the time). The tessellation divides the plane into polygons, each enclosing one tail position. Every polygon comprises the locations on the plane, which are closer to the position of the head enclosed by that polygon than to any other head.

Moreover, to quantify whether and how many lipids undergo face transition during the simulations, the regular packing of each of their chains was quantified by the hexagonal order parameter S_6, as previously reported in the literature.[88] Specifically, a chain was represented by its position on the xy plane (parallel to the membrane surface), computed as the average x and y positions of its carbon atoms. For each chain j, the set of neighboring chains was defined as the ones within a 0.65 nm radius from j. Then, S_6 is defined as

$$S_{6,j} = \frac{1}{6}\left|\sum_k e^{6i\theta_{jk}}\right| \tag{8}$$

with θ_{jk} being the angle between the vector connecting j and k and the x axis (i is the imaginary unit). A chain is in the gel phase if it has an hexagonal order parameter larger than 0.72.[88]

3. RESULTS AND DISCUSSION

In general, the parameters described in this work are shown to reproduce the available experimental target values well while, at the same time, are likely to allow for a better description of lipid−protein interactions, since the head groups are updated to the recent GROMOS force field.

3.1. Area per Lipid and Isothermal Area Compressibility Module. We report in Table 3 and Figure 3 the values of ApL for the simulation run. From such computations, it emerges that the increase of the CH3p−OM repulsion has a nonlinear effect on the area per lipid, as reported in ref 54. On the contrary, the comparison between simulation ID1 and the control ID0 for DPPC, which differ only in their partial charges, shows an almost identical ApL value (SI Figure 4). This suggests that the charge redistribution in the head group affects the global structure of the bilayer and the ApL less dramatically than the adopted value for the Lennard-Jones repulsion.

The comparison with the control simulation using a smaller membrane shows that larger systems allow for the evaluation of the ApL with a smaller error, as local fluctuations are averaged over a larger number of lipids. The standard deviation computed for the 128 lipids system is compatible with those reported in both the original[61] and a more recent publication,[89] in which the same system size was used.

The ApL from the simulation with a reaction field treatment for the long-range electrostatic term does not differ significantly from the one obtained with a PME treatment, being only slightly higher, which is in consistence with what was found in ref 90.

Finally, the values found using the Chiu charge set and the 54A7 force field (ID1 in Table 1) are systematically lower than those obtained in the original publications,[44,61] despite employing the same charge set and force field, while a better agreement is shown with those obtained more recently by Reif et al. for DPPC.[54] We attribute this to the different versions of GROMACS used, as the integration algorithm has recently been updated, affecting the calculated properties. Moreover, the double-cutoff scheme is no longer supported, preventing a faithful reproduction of the simulation setup used in ref 61. The variability caused by these changes has been extensively investigated by Reißer et al.[89] and reflects the observed discrepancy between the present and previous results.

From the considerations above, we suggest parameter set RM/54A8_v1 as the one that best reproduces all of the tested lipids at once. For DOPC and POPC bilayers, however, parameter set RM/54A8_v2 performs slightly better: it must be noticed that these two species present unsaturated bonds along the tails, whose influence might not be fully represented by any of the parameter sets. Indeed, it has been suggested that only a polarizable force field would be able to correctly capture the dynamics of the hydrophobic region of the membranes,[91]

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

Journal of Chemical Theory and Computation Article

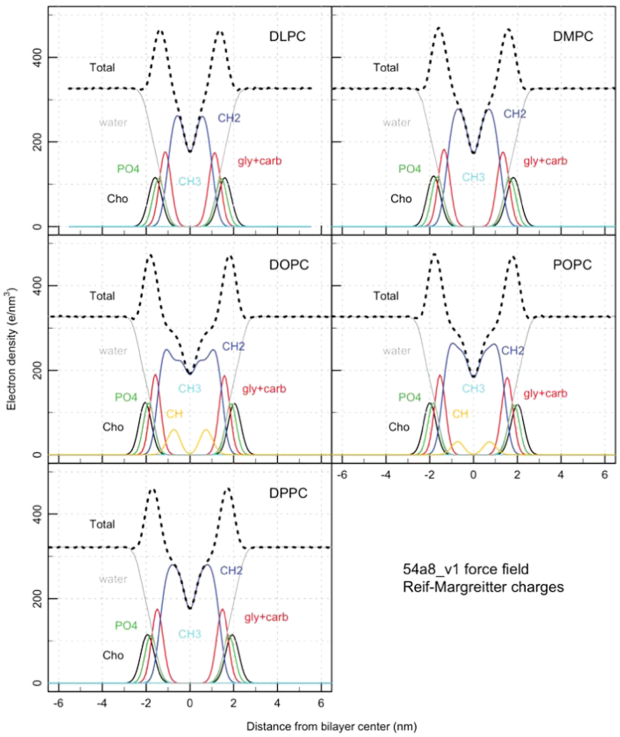


Figure 4. Electron density profiles of the hydrated DLPC, DMPC, DOPC, DPPC, and POPC bilayers (total) and of their individual components (Cho: choline, PO4: phosphate, gly + carb: glycerol and carbonyl groups, CH2: methylenes of the acyl chains, CH: CHdCH groups in the oleoyl chains, CH3: terminal methyls of the acyl chains) for simulation ID3 (54A8_v1 force field, Reif−Margreitter charges).

Table 4. Bilayer Thickness for Phosphocholine Bilayers, Derived from the Electron Density Profiles (Example in Figure 4) According to the Phosphate or Luzzati Methods[a]

hydrophobic thickness D_{HH} (nm)						
ID	charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	2.83	3.59	3.05	3.30	4.30
3	RM/54A8_v1	2.72	3.48	2.89	3.22	4.06
experiment[b]		3.08	3.44−3.60	3.53−3.71	3.70	3.42−3.83
Luzzati thickness D_B (nm)						
ID	charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	3.11	3.54	4.13	4.00	3.93
3	RM/54A8_v1	3.04	3.48	3.94	3.88	3.75
experiment[b]		3.14	3.63−3.96	3.59−3.87	3.68	3.50−3.83

[a]All simulations were run at 303 K, except for DPPC (323 K). [b]Values from ref 44 and Table 2 in ref 61.

taking in proper account the difference between saturated and unsaturated bonds.

For POPE and POPG, we resorted to the modification proposed by Kukol,[46] i.e., the use of the CH0 atom type for the ester carbons (see Section 2). For both lipids, a good agreement with experimental ApL values could be achieved using the new partial charge parameters (Figure 3).

Along the same lines, when comparing the results with the ones obtained with the CHARMM36 force field in its original publication,[65] we find DOPC, presenting an unsaturated bond in each tail, to be the most diverging. In particular, CHARMM36 better captures the spacing between the lipids, enhanced due to the presence of the double bond, and we suspect that this is due to its all-atom description.

Results of the isothermal area compressibility calculations confirm the finding of refs 61 and 65 that K_A values obtained from simulation are about 1.5−3 times larger than those measured experimentally (SI Table 3). This holds for all parameter sets tested. Set RM/54A8_v1 performs better than

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

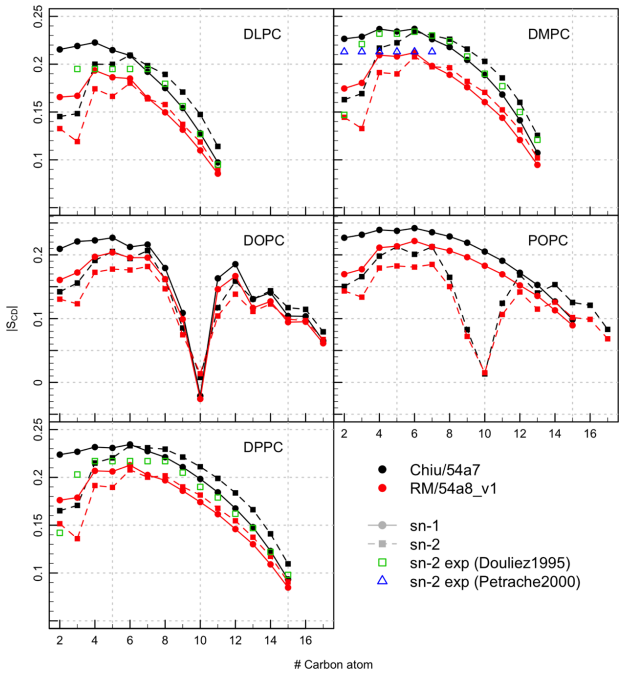


Figure 5. Deuterium order parameter S_{CD} profiles of the *sn*-1 (solid curves) and *sn*-2 (dashed curves) fatty acyl chains of hydrated DLPC, DMPC, DOPC, DPPC, and POPC bilayers calculated from simulations ID1 (54A7 force field, Chiu charges) and ID3 (54A8_v1 force field, Reif–Margreitter charges). The S_{CD} values are averaged over all of the lipid *sn*-1 and -2 acyl chains in the systems (proS hydrogen only). Experimental values Douliez1995 from ref 94 and Petrache2000 ones from ref 95.

Chiu/54A7 for all of the lipids tested but DLPC, for which the results are equivalent.

The overestimation of the compressibility is likely due to the underestimation of ApL fluctuations during dynamic simulations. The K_A value computed for the small, 128 lipids, DPPC bilayer patch is smaller than the one computed for the 512 lipid ones (342 and 499 mN m^{-1}, respectively), as the small patch exhibits higher fluctuations of the ApL (see Section 3.1). It is thus evident that the size of the system plays a pivotal role in obtaining correct fluctuations and global properties.

3.2. Electron and Charge Density Profile. Across simulations with different parameters, the electron density qualitatively maintains the same profile for each phosphocholine lipid. In Figure 4, the density for parameter set 54A8_v1 and all of the lipids is displayed (SI Figures 10–13 show the same plot for the other parameter sets), while in SI Figures 5–9, panel (b), the total and the phosphate group electron densities are shown for the five parameter sets tested, for one lipid at the time. The peak broadness shows a direct relationship with the packing density of the bilayer: larger ApL values correspond to a shallower profile of the density, due to fluctuations of the membrane along the *z* axis and to deeper penetration of water molecules into the bilayer.

The bilayer thickness was evaluated from the electron density profiles, as explained in Section 2. Our parameters are overall in better agreement with the Luzzati estimate of the thickness rather than the hydrophobic one, but altogether, these measurements (phosphate and Luzzati thickness) do not strongly discriminate between sets. In Table 4, the values for the Chiu/54A7 and RM/54A8_v1 sets are shown (see SI Table 4 for the complete results).

Further comparison of the dipole potential profiles, obtained from the charge density, shows how the RM/54A8_v1 charge set gives results closer to the ones obtained in all-atom simulations[82,92] (see SI Section 1 for a complete discussion).

3.3. Order Parameter of the Acyl Chains. For all of the lipids and parameter sets, S_{CD} is lower than 0.25, which indicates that the tails are generally disordered and the membrane has not transitioned to a gel-like state,[93] even for the simulation with the lowest ApL. Figure 5 and SI Figures 14–17 display the computed values for specific parameter sets, and SI Figures 5–9, panel (c), show a cross-parameter comparison for each lipid. Comparing these different sets, simulations denoted by ID from 1 to 5 show a consistently decreasing S_{CD}, in line with the increased area per lipid and decreased bilayer thickness. This indicates that when the lipid molecules are constrained in space, their tails tend to be stretched and ordered. The presence of unsaturated bonds in the DOPC and POPC lipids is captured, by all parameter sets, as a decrease in S_{CD} at the positions related to those bonds. The main difference due to the introduction of the new charges is in the decreased order observed for the first and second carbon bonds of the *sn*-1 tail, which show S_{CD} values smaller than the ones for the third carbon bond, while with the Chiu charges, a constant increase is observed with decreasing carbon index for tail *sn*-1.

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175–5193

Overall, the RM/54A8_v1 set is within the range of experimental values[94,95] (Figure 5); in particular, it captures the low order of the first *sn*-2 carbon atom (numbered 2) well, while the Chiu/54A7 set presents closer values in the central region of the tails. However, it must be noticed that variability is found within the experimental data (see the different experimental values reported in Figure 5). Therefore, without aiming at a perfect fit to such a small pool of experimental data, we consider set RM/54A8_v1 as sufficiently accurate in representing the experimental findings, in particular, in better reproducing the regions in the vicinity of the head group, while the description of the hydrophobic core remains less accurate and subject to improvement.

3.4. Hydration of Head Groups and Glycerol/Carbonyl Moieties. The hydration of functional groups of lipids is a key characteristic for both their dynamics and potential interactions with other molecules, such as proteins. From the distribution of distances between water oxygens and the nearest atom of various lipid groups, it emerges that the new partial charges modify the hydration profile of the lipid head group (Figure 6 shows the comparison between parameter sets for DPPC and SI Figures 17−20 for the other lipid bilayers).

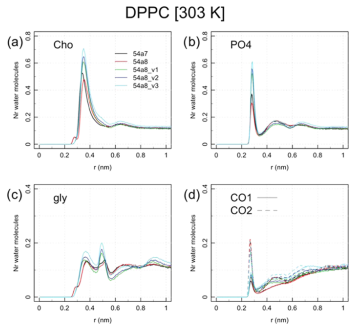


Figure 6. Distribution of the distance between the water oxygen and the nearest lipid head group atom for simulation DPPC. Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the *sn*-1 and *sn*-2 positions.

The choline major peak at 0.38 nm and the phosphate one at 0.30 nm are higher and sharper when employing the RM charges rather than the Chiu ones, reflecting an increased average hydration of these two moieties. Additionally, for the simulations run with the RM charges, the choline profile does not display the first, low intensity, peak obtained with the Chiu set at 0.28 nm: indeed, the charges of choline and the modification of the C12 Lennard-Jones repulsion for the NL atom type introduced in parameter set 54A8 were optimized to successfully prevent oversolvation, repelling water from its core.[47,54] The profiles of the other components are partially influenced, as well. For the RM charges, the second peak for glycerol increases its value and the two ester peaks have more similar values between them (Figure 6, panel (c), and SI Figures 17−20), which is consistent with deeper water penetration.

To quantify the observed differences, the hydration profiles were integrated up to the first peak or the second one in the case of phosphate and glycerol (SI Table 5). The results show that the average number of water molecules around the choline group is higher for the RM/54A8_v1 set than for the Chiu/54A7 one by one water molecule. This seems to contrast with the increased hydrophobicity of the newly parameterized choline moiety; however, this might partially be explained by the changed orientation of the head groups (see Section 3.5) and by the new parameterization of phosphate,[48] which accounted for the hydrogen bond potential of the most solvent-accessible atoms, leading to a better solvation of the head group in comparison to the Chiu/54A7 and Chiu/54A8 sets.

The integration up to the second peak of the distribution of distances between the water oxygens and any lipid head group atom gives values between 12 and 17 water molecules per lipid, which is in agreement with the experimental range of 10−20.[96−99] Again, parameter set RM/54A8_v1 results in more hydrated head groups (about one water molecule more for each lipid) with respect to Chiu/54A7. Notably, the average number of water molecules increases, as expected, for the simulations resulting in a larger ApL (RM/54A8_v2 and RM/54A8_v3). The trends above are confirmed by solvent-accessible surface area values, which are higher for the choline head groups described by the RM charge set with respect to the Chiu one, while the values are closer between parameter sets for the phosphate and glycerol moieties, which are more deeply buried (SI Figure 21).

The increased hydration might be of relevance when simulating interactions with peptides and proteins. Moreover, as shown in a recent comparison between different lipid force fields,[65] the Chiu/54A7 parameter set results in a slightly less hydrated head group with respect to the CHARMM36[37] and Lipid14[101] force fields; therefore, the new set of parameters achieves values closer to them.

3.5. Orientation of the Head Groups and Carbonyl Moieties. The orientation of the head groups, defined by the angle of the P−N vector with the outward bilayer normal, is similar for all of the lipids within the same parameter set (see SI Figure 22, top row). This indicates that the nature of the tails does not strongly affect the behavior of the head group, which is to be expected. Comparing different sets for DPPC

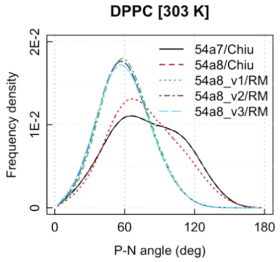


Figure 7. Distribution of the P−N, CO1, and CO2 angles with respect to the outward normal to the bilayer.

(Figure 7 and SI Figure 23), it emerges that with the Chiu charges, the distribution of P−N angles is bimodal, with preferred values around 60 and 90°, while the new charges restrict the motion to the 60° configuration. Recent experimental data support a value around 60° (see refs 100 and 102), as opposed to 90° as reported previously.[103]

It is noteworthy that this property was not part of the calibration process, i.e., the agreement with the experimental

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

observables in ref 102 is most likely due to a more accurate description of the solvation of the choline and phosphate moieties. Simulations performed by Botan et al.[104] confirm that smaller angles with respect to the membrane normal are caused by a higher level of head group hydration, which is in line with conclusions from the previous section. This difference in the predominant configuration of the lipids' head group will most probably influence the interaction with proteins or peptides approaching the interfacial region, providing a different binding recognition landscape.

The orientation of the *sn*-1 and *sn*-2 carbonyl dipoles with respect to the bilayer normal is again similar across different lipids (SI Figure 22, middle and bottom rows). The introduction of the RM charges has a small effect on these dipoles, as a result of the spatial rearrangement of the nearby head group. The most probable value for CO1 is shifted from 110 to 120° (Chiu vs RM charges), while the one for CO2 from 135 to 150°.

3.6. Lipid Lateral Diffusion. To correctly reproduce the membrane and its functions, its dynamical characteristics are as important as its structural ones. To address this, lipid lateral diffusion can be measured and compared against experimental data. Lateral diffusion is influenced by the area per lipid, with a tighter packing preventing larger displacements but is not solely determined by it.

Lateral diffusion coefficients (D) measured from simulations are shown in Figure 8. As anticipated, the set with largest ApL

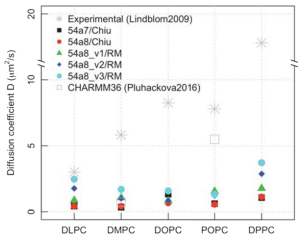


Figure 8. Lateral diffusion coefficient of DLPC, DMPC, DOPC, DPPC, and POPC bilayers for different parameter sets.

(54A8_v3) presents the highest values; however, parameter set 54A8_v1 gives significantly higher diffusion coefficients than those obtained with the Chiu/54A7 set, despite the values of ApL being similar. SI Figure 24 depicts a comparison of the diffusion coefficient of DPPC between ID0 and ID1, which differ only in the partial charges of the head groups. It confirms that the RM charges (ID0) allow for more mobility of the lipids with respect to Chiu ones (ID1), independent from all other modifications to the force field.

Regarding the simulation conditions, the use of a reaction field scheme increases the mobility by 34%, whereas the size of the patch decreases it by a small but significant amount (19%; see SI Figure 24). It is known that periodic boundary conditions affect the evaluation of lipid diffusion;[105,106] therefore, the larger the system simulated, the more accurate the reproduction of the experimental values. However, the change in the D due to the electrostatic treatment and the patch size, taken in absolute terms (i.e., a difference of about 0.4 and 0.2 $\mu m^2\ s^{-1}$, respectively), are small in comparison with the effect due to the adoption of the new parameters (between 2 and 6 $\mu m^2\ s^{-1}$).

The comparison with experimental values is challenging due to the fact that different experimental techniques report values, which are an order of magnitude apart. Poger et al. gave an overview of this variability for DPPC bilayers in Table 2 of ref 77 and observed that values span from 0.5 to 50 $\mu m^2\ s^{-1}$. In this view, the values obtained in the present work for DPPC are well within the range, regardless of the parameter set chosen. However, we report experimental values from ref 107 obtained through pulsed-field gradient nuclear magnetic resonance as a guide. Additionally, we report the values obtained with CHARMM36 in ref 65. The CHARMM36 benchmarks are present only for two of the phosphocholines analyzed in this work and show that the values obtained with this force field span a broader range. The consistently low values of D computed with the different GROMOS parameter sets in this work are in agreement with what was found in the literature.[108,109]

3.7. Tilt Modulus. We report in Figure 9 the values of κ_t^1 obtained for each of the phosphocholines considered and each

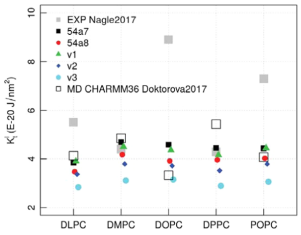


Figure 9. Tilt modulus κ_t^1 computed from the distribution of lipid tilt angles along the last 300 ns of the trajectories. The results are compared with the experimental values from ref 110 and the ones obtained (with the same procedure as employed here) from CHARMM36 simulations in ref 111.

parameter set tested. For comparison, we plot the experimental values obtained by Nagle et al. in ref 110 and the results from simulations using the CHARMM36 force field.[111] Given that the data show quite a large spread in their values depending on the actual experimental setup used, for the comparison, we selected values, which were all obtained under the same conditions, for both the experiments and the computational results.

The plot shows that the tilt modulus κ_t^1 varies between 3 and 5×10^{-20} J nm^{-2}. In simulations resulting in larger ApL (e.g., parameter set 54A8_v3), the lipids are in a less dense environment and can better accommodate changes in their orientations resulting in a lower tilt modulus (the tilt modulus gives the energy necessary for tilting the lipids per unit area).

The comparison with the experimental values is very good for DMPC and DPPC, while it is poorer for DLPC and very poor for DOPC and POPC, which harbor unsaturated bonds in the tails. Results from the CHARMM36 simulations show more variability between different phosphocholines, but a similar if not lower agreement with the experiment (for example, for DOPC). In general, comparing the results together, we think that we achieved a sufficiently qualitative agreement with the previous computational literature.

The discrepancy with experiments (both for our results and for the ones from ref 111) is likely due to computational limitations: as the tilt is retrieved from an ensemble

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

distribution, larger and longer simulations are more likely to give a better result. Moreover, as briefly mentioned at various points in the manuscript, artifacts arising in the hydrophobic regions of lipids (such as the suboptimal modeling of unsaturated tails) will probably only be resolved using a polarizable description.

3.8. Phase Transition Behavior. The previous analysis points to parameter set 54A8_v1 as the one that best reproduces the experimental properties for each of the lipids simulated. Therefore, we test this set further to assess its ability to reproduce the change in lipid behavior under different temperature conditions. As mentioned in Section 2.6, we use as test system the DPPC bilayer patch.

The comparison between simulations at 323 and 333 K shows that the global area per lipid increases with temperature, consistently with what was expected. Parameter set 54A8_v1 captures the increase in lipid spacing, with a slight underestimation of ApL at 333 K with respect to the experiments (for 54A8_v1 and experiments, respectively: 0.624(6) vs 0.631(13) nm^2 at 323 K and 0.634(5) vs 0.650(13) nm^2 at 333 K). The experimental ApL is measured at a different temperature with the same experimental setup.[16]

The simulation at 303 K shows the formation of a patch of ordered lipids, suggesting that the parameters can reproduce different phases: two nucleation sites for the gelification process are observed in both the upper leaflet and lower leaflet, in nonmatching positions, and the gel front extends over time. To classify the phase a lipid belongs to, we computed the hexagonal order parameter S_6 for each chain from the last frame of the simulation (time point 400 ns). Figure 10 shows

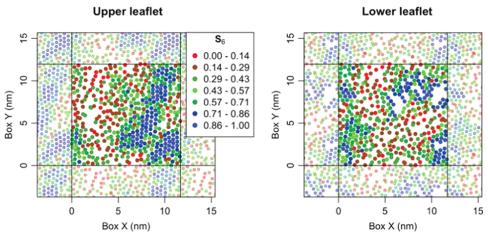


Figure 10. Hexagonal order parameter S_6 for lipid acyl chains computed on the last frame of a DPPC bilayer simulation using 54A8_v1 parameters. Each point corresponds to the average position of the carbon atoms of the respective chain on the *xy* plane. Thus, every lipid is represented by two points in these plots. Solid black lines denote the boundaries of the simulation box and chains of the periodic images (used for the computation of S_6 boundaries) are shown faded out. Colors from red to blue denote an increasing S_6 value: the last two indicate gel-phase lipids.

the position of the chains on the *xy* plane, for each leaflet separately, color-coded by S_6: the regions where S_6 is larger correspond to a densely packed area with a quasi-hexagonal lattice. In particular, the center of the ordered patches has S_6 values larger than 0.72 (last two colors of the scale), i.e., it can be classified as a gel. Overall, 20% of chains have undergone this transition within the time simulated.

Averaging over all of the lipids, S_6 at 303 K is 0.45. As a comparison, we computed this average quantity on the last frame of the simulations performed at 323 and 333 K, finding 0.28 and 0.27, respectively (with only six and three chains above the gel threshold of 0.72).

A hexagonal order can be obtained when the tails are well ordered and parallel to each other, standing in a vertical straight conformation (SI Figure 25 shows a detail of a well-ordered gel patch). We thus compute the S_{CD} order parameter of the acyl chains averaging separately over the lipids for which at least one chain has an S_6 value larger than the 0.72 threshold (168 lipids overall) and for the others. The last 100 ns of the simulation time was used. These values are compared to the average S_{CD} from the simulations at 323 and 333 K. Figure 11

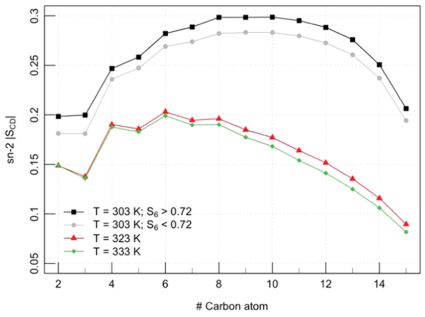


Figure 11. Order parameter for the acyl chain *sn*-2 for a DPPC bilayer simulated at a different temperature. The average is performed including both leaflets. For the simulation at 303 K, the lipids were split in two groups according to their hexagonal order parameter S_6 and the acyl chain order parameter S_{CD} computed for each of them.

shows highly ordered tails for the membrane simulated below the transition temperature, for both the gel and nongel lipids (classified according to the S_6 threshold). This suggests that the full patch is undergoing a phase transition, but the completion of the process is not seen due to the short simulation time scale. As a comparison, tails at 323 and 333 K are much more disordered, with a slight decrease in order with increasing temperature.

Finally, we computed the local area per lipid chain from a Dirichlet tessellation of the same set of points used to calculate the hexagonal order parameter. The average values over the gel-phase tails (multiplied by 2) give an ApL of 0.438 ± 0.038 nm^2, while the remaining of the chains have widely spread values, correlated to their S_6 parameter, giving an average of 0.57 ± 0.18 nm^2. The values found for the gel patches (at 303 K) are close to the experimental outcomes by Nagle et al. of 0.473 nm^2 at 293 K[112] and 0.479(2) at 297 K.[113] The value computed from simulations is smaller likely because it is computed only over the tails perfectly packed in a hexagonal lattice.

Altogether, these results prove that the newly developed parameters can successfully reproduce the gel phase when a lipid patch is simulated below the phase transition temperature.

3.9. Transferability to POPE and POPG. As mentioned above, the areas per lipid values of POPE and POPG simulations, where the phosphate partial charges have been replaced with the RM values and, in the case of POPE, the amine partial charges have been updated according to the 54A8 force field, are in good agreement with the available experimental data.

For POPE, a slightly enhanced hydration is obtained from the update of the phosphate charges (from 5.7 to 6.4 water

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

Journal of Chemical Theory and Computation Article

Table 5. Binding Time of Lactoferricin (LFC) Peptide to the Model Membranes in Examination[a]

	binding time (ns)							
	DLPC/DLPG 3:1				POPC			
	OI	OII	OIII	OIV	OI	OII	OIII	OIV
Chiu/54A8	0.5	3.8	6.3	2.8	13.9	21.9	(65.0)	13.9
RM/54A8_v1	2.8	75.2	62.0	9.6	1.0	NA	NA	NA

[a]NA denotes no binding observed in the time simulated (100 ns). For POPC/OIII simulated with parameter sets Chiu/54A8, the binding time is in parentheses as LFC approaches the membrane but maintains a 0.4 nm distance (±0.02), which is higher than the threshold chosen to define a binding event.

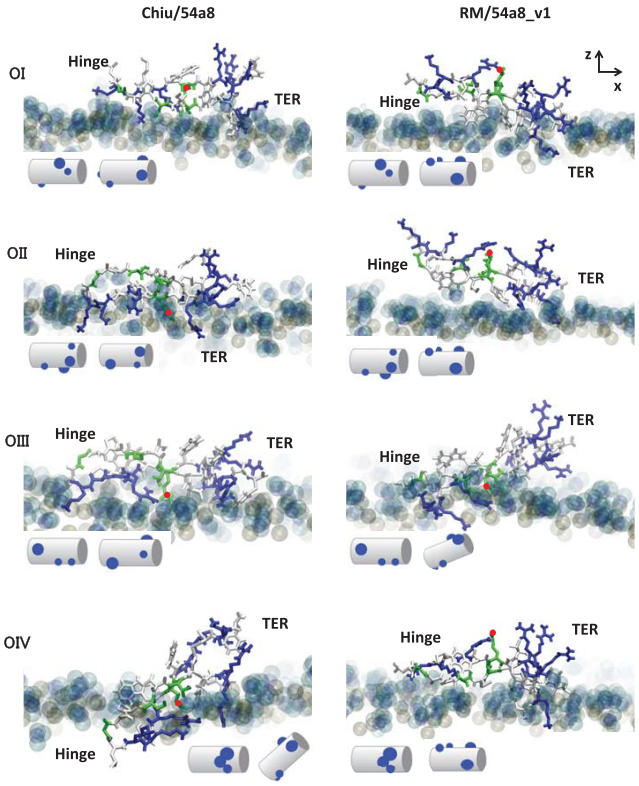


Figure 12. Final configurations of the simulations of LFC on a DLPC/DLPG 3:1 membrane, starting from four different initial orientations OI–OIV. OII, OIII, and OIV are obtained from OI with an anticlockwise rotation of, respectively, 90, 180, and 270° along the main axis (*x* axis in the top right panel). LFC is colored by residue type (blue charged, green hydrophilic, white hydrophobic); phosphorus atoms are shown in golden beads. The terminal and hinge regions of 1LFC are indicated (TER, hinge), together with GLN7 as a red dot to help the visualization. The insets shows a cartoon representation of the initial and final configurations, highlighting the positive patches as blue dots.

molecules per lipid with experimental values between 4 and 7;[114,115] see SI Table 7) with similar results in terms of thickness D_B (3.89 nm for Chiu and 3.92 for RM set, with an experimental value of 4.13 nm;[116] SI Table 6 and SI Figures 26 and 27). Overall, these results confirm the transferability of the new phosphate charges to different types of phospholipids.

4. INTERACTION WITH PROTEINS

The adoption of the updated parameters enhances the consistency with the GROMOS parameters for protein simulations. To test how this affects the simulations of peptides interacting with a membrane, we performed additional simulations of a small antimicrobial peptide on the surface of two different model membranes.

The peptide selected is bovine lactoferricin (PDB code 1LFC). It has a length of 25 amino acids and adopts a β-hairpin conformation in solution, with many aromatic hydrophobic residues on one side and charged amino acids distributed all over.[117] This peptide is antimicrobial and therefore found to preferentially bind bacterial membranes

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

versus mammal ones, as shown by NMR experiments on LFC subsequences (namely, LFC_{4-9}[118] and LFC_{4-14}[119]).

One can model the bacterial membrane by a mixture of zwitterionic and anionic lipids. The latter are characteristic of the cell wall of both Gram-positive and -negative bacteria.[120,121] In this study, we selected the mixture DLPC/DLPG with a 3:1 ratio that has been used to elucidate the antimicrobial activity of lactoferricin-derived peptides.[122] As for the mammal membrane description, we used POPC as it has been often used in molecular dynamics simulations with this purpose.[123−125] Despite being rather simple, these or similar model membranes have often been used in experiments to test, among others, the effects of antimicrobial peptides upon binding.[14]

Molecular simulations can shed light on the differences in the binding process of LFC to antimicrobial and mammal model membranes. Our parameterization should then reflect a sensible difference in the binding behavior for these two cases.

The exact binding mechanism of LFC to a membrane is not fully understood, and a number of experimental papers have hypothesized binding modes for the interactions of this peptide with model membranes. A mutation study in LFC_{1-15} suggests that Trp residues anchor the peptide to the membrane as the antimicrobial activity of the peptide was retained only when Trp was mutated in equally hydrophobic amino acids.[126] However, the role of Trp seems to be different in other antimicrobial peptides, where they reside at the lipid−water interface and form hydrogen bonds with the moieties nearby.[127] As experiments on the full-length peptide (25 amino acids) have not yet been reported and the full sequence contains additional charged and hydrophobic residues, it remains unclear whether this additional region would change the aforementioned binding mechanism. We therefore decided to use molecular dynamics simulations to elucidate molecular determinants in discriminating the binding of the peptide to mammal and bacterial membranes.

In order to not be biased by the initial configuration adopted in the simulation, we performed multiple simulations with different initial orientations of the peptide relative to the membrane. The hairpin main axis was aligned to the membrane plane and the peptide rotated around this axis in steps of 90°, leading to four different starting orientations named OI, OII, OIII, and OIV (SI Figure 29). This allows different segments of the sequence (and thus amino acids with different chemical characteristics) to face the membrane in the initial positioning. The initial minimum distance between the peptide and the lipids was set to 2 nm. The simulation length was 100 ns each, sufficient to see the binding process in all of the control cases.

The simulations have been performed for the proposed RM/54A8_v1 parameter set and the Chiu/54A8 one (control cases) to compare with the most recent set available in GROMOS and highlight the difference of the newly parameterized lipid head groups.

To quantify the outcome of the simulations, we monitored the time at which the peptide binds (always irreversibly) to the membrane as the time at which the minimum distance between the peptide and the membrane is below 0.3 nm (Table 5). The cutoff was chosen, analyzing the configurations after LFC bound to the membrane, which resulted generally to stabilize around a minimum distance of 0.25 nm. The minimum distance was computed every 100 ps, and a running average was applied with a 10 frame window. Additionally, the insertion depth of each amino acid in the membrane has been calculated as the difference between the z position of the lowest atom of the amino acid and the average of the maximum z coordinate of the five lipids closest to it.

Table 5 shows the different binding times for LFC against a mixed DLPC/DLPG or pure POPC membrane patch. For the mixed, anionic membrane, the new parameters favor a slower and weaker binding process. Indeed, with parameter set Chiu/54A8, the peptide is quickly sequestrated by the lipids due to the opposite charge interaction. This favors an unspecific binding, dependent on the sequence facing the membrane in the initial configuration (Figure 12 and SI Movie 1). In Figure 13, the average insertion in the membrane after the binding is

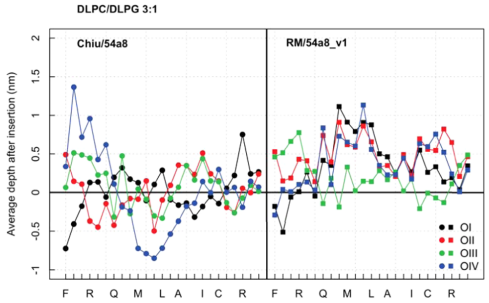


Figure 13. Average insertion depth of each amino acid after binding of the LFC peptide as per Table 5. The zero value is the top of the membrane plane so that a negative depth means insertion into the membrane. Some reference amino acids are displayed at the bottom.

plotted for each amino acid: Chiu/54A8 favors a deep insertion of differently charged residues for different runs. The RM/54A8_v1 simulations produce a less inserted configuration of LFC and a more consistent protrusion of the hinge region out of the membrane, i.e., the central stretch of amino acids between Met and Leu (Figure 13), as three out of the four simulations (all but OIII) show this behavior.

The angular orientation of the peptide around its axis has been computed as the angle formed with the z axis by the backbone carbon and nitrogen bonded via a hydrogen bond (amino acids 7 with 19 and 9 with 17), confirming that the new set of parameters allows for more freedom in the reorientation of the initial configurations, while the previous one tends to keep them close to the original configuration (SI Figure 30). Additionally, the new set of parameters seems to favor the reorientation of the peptide as to face the Trp residues toward the membrane surface in three out of the four simulations, in contrast with the results from the previous parameterization. This preference for the interfacial region is a known mechanism in the membrane binding of Trp- and Arg-rich peptides.[127]

When simulating a pure POPC membrane (here considered as a mammal membrane model), the resulting binding poses obtained with the Chiu/54A8 and different initial conditions are consistent among each other; in particular, the three amino acids Lys12−Leu13−Gly14 located at the hinge of the hairpin promote the insertion, while the terminal region stays exposed in solution. Therefore, parameter set Chiu/54A8 discriminates between the two membranes as it suggests a weaker binding to the mammal one. However, with the parameter set RM/

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

54A8_v1, three out of the four simulations result in no binding at all, in agreement with experimental findings.[118] The remaining simulation (OI) shows a quick binding event, promoted by the terminal regions as observed for three out of four simulations with the DLPC/DLPG 3:1 membrane.

The results above highlight that the new parameters show a membrane-binding process less dependent on the initial conditions, allowing for a dynamical rearrangement of the protein at the membrane interface. This comes at the expenses of a longer sampling time needed to observe binding events for most of the configurations chosen. Future work will focus on systematic comparisons of available peptide−membrane simulations with other parameterizations and on longer simulated times.

The difference between the behavior on a model bacterial or mammal membrane is more pronounced for the new parameters, and this is consistent with the selective antimicrobial action of the peptide and its low hemolytic activity.[118,119] Overall, we think that these new parameters show promising characteristics for the simulation of membrane−peptide interactions within the GROMOS force field, particularly for the study of interfacial absorption.

5. CONCLUSIONS

In this work, we present a reparameterization of a range of phospholipids in the context of the GROMOS force field, taking advantage of recent optimizations reported for key chemical groups in these molecules. The effect of the newly adopted head group partial charges has been tested extensively to ensure that they match experimentally observable characteristics of lipid bilayers. In parallel, we tested the effect of the van der Waals repulsion between the choline methyl groups and the phosphate oxygens, as it was modified by Poger et al. to reproduce the experimental area per lipid values while using the partial charges derived by Chiu et al. A summary of the updated parameters and simulation conditions is available in SI Table 10.

The work proves that the new charges are suitable to describe all of the phosphocholine bilayers tested, matching the experimental values as successfully as the previous parameter set. The major advantage of the Reif−Margreitter set lies in the partial charges of the head group, which are derived by applying the GROMOS parameterization philosophy rather than quantum mechanics calculations, thereby providing a description, which is more consistent with the parameters adopted for other biomolecules such as proteins within this force field. By using the updated partial charges for the choline (more hydrophobic) and phosphate (more hydrophilic) groups, the parameters also show a better reproduction of the average head group orientation, which was recently reassessed by experiments. The value of the Lennard-Jones repulsion term found to best reproduce the experimental values is the one in set 54A8_v1, which is set to a value in between that of the 54A7 and 54A8 parameter sets.

In the Reif−Margreitter parameter set, only the partial charges of the ester groups remain as described in the Chiu charge set. Preliminary work has been started to test the influence of the ester charges in combination with the new ones for the head group but, in accordance to what was previously found by Chandrasekhar et al. (ref 59), the replacement of the ester charges with the standard ones for the ester moiety resulted in values of area per lipid too low with respect to the experimental findings. As mentioned previously, it is possible that this discrepancy with the rest of the force field can only be avoided by adopting a polarizable force field.[91] However, in the absence of further sophisticated changes to the force field parameterization, we are confident that the proposed parameters are a major step forward in the description of lipid head groups and they should enable improved modeling of the interaction of lipids with water and other soluble molecules.

The new phosphate partial charges have been proved to transfer well to other phospholipids not presenting a choline head. For those lipids, the Kukol modification, which takes advantage of a different atom type for the ester carbon, is adopted to obtain the correct area per lipid.

Finally, the performance in reproducing some specific peptide−membrane interactions was tested. In this respect, the new parameter set shows significant differences with respect to the latest Chiu/54A8 set: it better discriminates the binding of an antimicrobial sequence on a bacterial versus a mammal membrane. Additionally, it favors a weaker and more dynamic binding, which is less biased from the initial conditions of the simulations.

In conclusion, we believe that the new Reif−Margreitter charge set together with the GROMOS 54A8_v1 parameter set is a major improvement on the previous iteration of the GROMOS lipid force field and should be particularly suited for protein−membrane systems, such as studies including small antimicrobial peptides, which rely on an accurate peptide−membrane recognition.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.9b00509.

Dipole potential computation

Additional tables: control simulations; non phosphocholine simulations; compressibility, thickness, hydration profile for all parameter sets; summary of simulation parameters

Additional plots: control simulations results; comparison of properties across parameter sets; electron density, order parameter, hydration profile and head group orientation for parameter sets not presented in the main text; phosphocholine moieties SASA; snapshot of DPPC patch partially in gel phase; non phosphocholine simulations results; scheme and measure of 1LFC orientation with respect to the membrane (PDF)

Simulation of LFC on a bacterial membrane for both parameter sets Chiu/54A8 (left half) and RM/54A8_v1 (right half), starting from orientation OII (Movie 1) (MP4)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: franca.fraternali@kcl.ac.uk.

ORCID

Irene Marzuoli: 0000-0001-7536-6144

Franca Fraternali: 0000-0002-3143-6574

Notes

The authors declare no competing financial interest.

The topology and the parameter files in GROMACS format for the lipid molecules and the atomic coordinates of equilibrated bilayers of DLPC, DMPC, DOPC, POPC, DPPC, POPE, and

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

POPG are available at http://fraternalilab.kcl.ac.uk/wordpress/biomembrane-simulations. Note that these parameter files include the post-translationally modified residues specified for the GROMOS 54A8 force field earlier.[128]

■ ACKNOWLEDGMENTS

The authors thank Chris Oostenbrink and Maria M. Reif for careful reading of the manuscript and feedback. I.M. acknowledges funding by the Engineering and Physical Sciences Research Council (EPSRC) through the Centre for Doctoral Training "Cross Disciplinary Approaches to Non-Equilibrium Systems" (CANES, Grant No. EP/L015854/1). C.M. acknowledges funding by the MRC/BBSRC Systems Immunology of the Lifecourse programme grant: MABRA "Multiscale analysis of B cell responses in ageing" (MR/L01257X/1). This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (http://www.csd3.cam.ac.uk/), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology Facilities Council (www.dirac.ac.uk).

■ REFERENCES

(1) Cooper, G. M. *Transport of Small Molecules*; Sinauer Associates, 2000.
(2) Cocucci, E.; Kim, J.; Bai, Y.; Pabla, N. Role of Passive Diffusion, Transporters, and Membrane Trafficking-Mediated Processes in Cellular Drug Transport. *Clin. Pharmacol. Ther.* **2017**, *101*, 121−129.
(3) Mälkiä, A.; Murtomäki, L.; Urtti, A.; Kontturi, K. Drug permeation in biomembranes: In vitro and in silico prediction and influence of physicochemical properties. *Eur. J. Pharm. Sci.* **2004**, *23*, 13−47.
(4) Scott, D.; Ghosh, A.; Di, L.; Maurer, T. Passive drug permeation through membranes and cellular distribution. *Pharmacol. Res.* **2017**, *117*, 94−102.
(5) Zhao, R.; Babani, S.; Gao, F.; Liu, L.; Goldman, I. D. The mechanism of transport of the multitargeted antifolate (MTA) and its cross-resistance pattern in cells with markedly impaired transport of methotrexate. *Clin. Cancer Res.* **2000**, *6*, 3687−3695.
(6) Shi, Y.; Massagué, J. Mechanisms of TGF-β Signaling from Cell Membrane to the Nucleus. *Cell* **2003**, *113*, 685−700.
(7) Jean-Louis, S.; Akare, S.; Ali, M. A.; Mash, E. A.; Meuillet, E.; Martinez, J. D. Deoxycholic acid induces intracellular signaling through membrane perturbations. *J. Biol. Chem.* **2006**, *281*, 14948−14960.
(8) Hodder, A. N.; Crewther, P. E.; Anders, R. F. Specificity of the protective antibody response to apical membrane antigen 1. *Infect. Immun.* **2001**, *69*, 3286−3294.
(9) Fahy, E.; Subramaniam, S.; Brown, H. A.; Glass, C. K.; Merrill, A. H.; Murphy, R. C.; Raetz, C. R. H.; Russell, D. W.; Seyama, Y.; Shaw, W.; Shimizu, T.; Spener, F.; van Meer, G.; VanNieuwenhze, M. S.; White, S. H.; Witztum, J. L.; Dennis, E. A. A comprehensive classification system for lipids. *J. Lipid Res.* **2005**, *46*, 839−862.
(10) Sharpe, H. J.; Stevens, T. J.; Munro, S. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell* **2010**, *142*, 158−169.
(11) Entova, S.; Billod, J.-M.; Swiecicki, J.-M.; Martín-Santamaría, S.; Imperiali, B. Insights into the key determinants of membrane protein topology enable the identification of new monotopic folds. *eLife* **2018**, *7*, No. e40889.
(12) Zhang, L.; Rozek, A.; Hancock, R. E. Interaction of cationic antimicrobial peptides with model membranes. *J. Biol. Chem.* **2001**, *276*, 35714−35722.
(13) Risselada, H. J.; Marrink, S. J. The molecular face of lipid rafts in model membranes. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17367−17372.
(14) Chan, Y.-H. M.; Boxer, S. G. Model membrane systems and their applications. *Curr. Opin. Chem. Biol.* **2007**, *11*, 581−587.
(15) Nagle, J. F.; Tristram-Nagle, S. Structure of lipid bilayers. *Biochim. Biophys. Acta* **2000**, *1469*, 159−195.
(16) Kučerka, N.; Heberle, F.; Pan, J.; Katsaras, J. Structural Significance of Lipid Diversity as Studied by Small Angle Neutron and X-ray Scattering. *Membranes* **2015**, *5*, 454−472.
(17) Pan, J.; Heberle, F. A.; Tristram-Nagle, S.; Szymanski, M.; Koepfinger, M.; Katsaras, J.; Kučerka, N. Molecular structures of fluid phase phosphatidylglycerol bilayers as determined by small angle neutron and X-ray scattering. *Biochim. Biophys. Acta, Biomembr.* **2012**, *1818*, 2135−2148.
(18) Kučerka, N.; Nieh, M.-P.; Katsaras, J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. Acta, Biomembr.* **2011**, *1808*, 2761−2771.
(19) Gurtovenko, A. A.; Vattulainen, I. Molecular Mechanism for Lipid Flip-Flops. *J. Phys. Chem. B* **2007**, *111*, 13554−13559.
(20) Tieleman, D.; Marrink, S.-J. Lipids Out of Equilibrium: Energetics of Desorption and Pore Mediated Flip-Flop. *J. Am. Chem. Soc.* **2006**, *128*, 12462−12467.
(21) Marrink, S.-J.; Mark, A. E. The Mechanism of Vesicle Fusion as Revealed by Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2003**, 11144.
(22) de Vries, A. H.; Mark, A. E.; Marrink, S.-J. Molecular Dynamics Simulation of the Spontaneous Formation of a Small DPPC Vesicle in Water in Atomistic Detail. *J. Am. Chem. Soc.* **2004**, *126*, 4488−4489.
(23) Marrink, S.-J.; Risselada, J.; Mark, A. E. Simulation of gel phase formation and melting in lipid bilayers using a coarse grained model. *Chem. Phys. Lipids* **2005**, *135*, 223−244.
(24) Marrink, S.-J.; Lindahl, E.; Edholm, O.; Mark, A. E. Simulation of the Spontaneous Aggregation of Phospholipids into Bilayers. *J. Am. Chem. Soc.* **2001**, *123*, 8638−8639.
(25) Skjevik, Å. A.; Madej, B. D.; Dickson, C. J.; Teigen, K.; Walker, R. C.; Gould, I. R. All-atom lipid bilayer self-assembly with the AMBER and CHARMM lipid force fields. *Chem. Commun.* **2015**, *51*, 4402−4405.
(26) Vernier, P. T.; Ziegler, M. J.; Sun, Y.; Chang, W. V.; Gundersen, M. A.; Tieleman, D. P. Nanopore Formation and Phosphatidylserine Externalization in a Phospholipid Bilayer at High Transmembrane Potential. *J. Am. Chem. Soc.* **2006**, *128*, 6288−6289.
(27) Leontiadou, H.; Mark, A. E.; Marrink, S.-J. Ion transport across transmembrane pores. *Biophys. J.* **2007**, *92*, 4209−4215.
(28) Marrink, S.-J.; Jähnig, F.; Berendsen, H. J. Proton transport across transient single-file water pores in a lipid membrane studied by molecular dynamics simulations. *Biophys. J.* **1996**, *71*, 632−647.
(29) Leontiadou, H.; Mark, A. E.; Marrink, S.-J. Molecular dynamics simulations of hydrophilic pores in lipid bilayers. *Biophys. J.* **2004**, *86*, 2156−2164.
(30) Sengupta, D.; Leontiadou, H.; Mark, A. E.; Marrink, S.-J. Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochim. Biophys. Acta, Biomembr.* **2008**, *1778*, 2308−2317.
(31) Leontiadou, H.; Mark, A. E.; Marrink, S.-J. Antimicrobial Peptides in Action. *J. Am. Chem. Soc.* **2006**, *128*, 12156−12161.
(32) Lipkin, R.; Pino-Angeles, A.; Lazaridis, T. Transmembrane Pore Structures of β-Hairpin Antimicrobial Peptides by All-Atom Simulations. *J. Phys. Chem. B* **2017**, *121*, 9126−9140.
(33) Piggot, T. J.; Holdbrook, D. A.; Khalid, S. Electroporation of the *E. coli* and *S. aureus* Membranes: Molecular Dynamics Simulations of Complex Bacterial Membranes. *J. Phys. Chem. B* **2011**, *115*, 13381−13388.
(34) Khalid, S.; Berglund, N. A.; Holdbrook, D. A.; Leung, Y. M.; Parkin, J. The membranes of Gram-negative bacteria: progress in molecular modelling and simulation. *Biochem. Soc. Trans.* **2015**, *43*, 162−167.

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

(35) Pino-Angeles, A.; Leveritt, J. M.; Lazaridis, T. Pore Structure and Synergy in Antimicrobial Peptides of the Magainin Family. *PLOS Comput. Biol.* **2016**, *12*, No. e1004570.
(36) Lopes, D.; Jakobtorweihen, S.; Nunes, C.; Sarmento, B.; Reis, S. Shedding light on the puzzle of drug-membrane interactions: Experimental techniques and molecular dynamics simulations. *Prog. Lipid Res.* **2017**, *65*, 24−44.
(37) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **2010**, *114*, 7830−7843.
(38) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D. Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *J. Chem. Theory Comput.* **2012**, *8*, 348−362.
(39) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71−73.
(40) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(41) Scott, W.; Hünenberger, P.; Tironi, I.; Mark, A.; Billeter, S.; Fennen, J.; Torda, A.; Huber, T.; Krüger, P.; van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.
(42) Chiu, S.-W.; Clark, M.; Balaji, V.; Subramaniam, S.; Scott, H. L.; Jakobsson, E. Incorporation of surface tension into molecular dynamics simulation of an interface: a fluid phase lipid bilayer membrane. *Biophys. J.* **1995**, *69*, 1230−1245.
(43) Tieleman, D. P.; Berendsen, H. J. C. Molecular dynamics simulations of a fully hydrated dipalmitoylphosphatidylcholine bilayer with different macroscopic boundary conditions and parameters. *J. Chem. Phys.* **1996**, *105*, 4871.
(44) Poger, D.; Van Gunsteren, W. F.; Mark, A. E. A new force field for simulating phosphatidylcholine bilayers. *J. Comput. Chem.* **2010**, *31*, 1117−1125.
(45) Chandrasekhar, I.; Kastenholz, M.; Lins, R. D.; Oostenbrink, C.; Schuler, L. D.; Tieleman, D. P.; van Gunsteren, W. F. A consistent potential energy parameter set for lipids: dipalmitoylphosphatidylcholine as a benchmark of the GROMOS96 45A3 force field. *Eur. Biophys. J.* **2003**, *32*, 67−77.
(46) Kukol, A. Lipid Models for United-Atom Molecular Dynamics Simulations of Proteins. *J. Chem. Theory Comput.* **2009**, *5*, 615−626.
(47) Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C. New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *J. Chem. Theory Comput.* **2012**, *8*, 3705−3723.
(48) Margreitter, C.; Reif, M. M.; Oostenbrink, C. Update on phosphate and charged post-translationally modified amino acid parameters in the GROMOS force field. *J. Comput. Chem.* **2017**, *38*, 714−720.
(49) Chandrasekhar, I.; van Gunsteren, W. F. Sensitivity of molecular dynamics simulations of lipids to the size of the ester carbon. *Curr. Sci.* **2001**, 1325−1327.
(50) Chandrasekhar, I.; van Gunsteren, W. F. A comparison of the potential energy parameters of aliphatic alkanes: molecular dynamics simulations of triacylglycerols in the alpha phase. *Eur. Biophys. J.* **2002**, *31*, 89−101.
(51) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843−856.
(52) Bransden, B. H.; Joachain, C. J. *Physics of Aoms and Molecules*; Prentice Hall, 2003; p 1114.
(53) Mulliken, R. S. Electronic Population Analysis on LCAO-MO Molecular Wave Functions. I. *J. Chem. Phys.* **1955**, *23*, 1833−1840.
(54) Reif, M. M.; Winger, M.; Oostenbrink, C. Testing of the GROMOS Force-Field Parameter Set 54A8: Structural Properties of Electrolyte Solutions, Lipid Bilayers, and Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 1247−1264.
(55) Van Lehn, R. C.; Alexander-Katz, A. Membrane-Embedded Nanoparticles Induce Lipid Rearrangements Similar to Those Exhibited by Biological Membrane Proteins. *J. Phys. Chem. B* **2014**, *118*, 12586−12598.
(56) Van Lehn, R. C.; Ricci, M.; Silvia, P. H. J.; Andreozzi, P.; Reguera, J.; Voïtchovsky, K.; Stellacci, F.; Alexander-Katz, A. Lipid tail protrusions mediate the insertion of nanoparticles into model cell membranes. *Nat. Commun.* **2014**, *5*, 4482.
(57) Jia, Z.; O'Mara, M. L.; Zuegg, J.; Cooper, M. A.; Mark, A. E. The Effect of Environment on the Recognition and Binding of Vancomycin to Native and Resistant Forms of Lipid II. *Biophys. J.* **2011**, *101*, 2684−2692.
(58) Berglund, N. A.; Piggot, T. J.; Jefferies, D.; Sessions, R. B.; Bond, P. J.; Khalid, S. Interaction of the Antimicrobial Peptide Polymyxin B1 with Both Membranes of *E. coli*: A Molecular Dynamics Study. *PLOS Comput. Biol.* **2015**, *11*, No. e1004180.
(59) Chandrasekhar, I.; Oostenbrink, C.; van Gunsteren, W. F. Simulating the Physiological Phase of Hydrated DPPC Bilayers: The Ester Moiety. *Soft Mater.* **2004**, *2*, 27−45.
(60) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656−1676.
(61) Poger, D.; Mark, A. E. On the Validation of Molecular Dynamics Simulations of Saturated and cis -Monounsaturated Phosphatidylcholine Lipid Bilayers: A Comparison with Experiment. *J. Chem. Theory Comput.* **2010**, *6*, 325−336.
(62) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **1995**, *102*, 5451−5459.
(63) Abraham, M. J.; van der Spoel, D.; Lindahl, E.; Hess, B. *GROMACS User Manual*, version 2016; GROMACS Development Team, 2018.
(64) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(65) Pluhackova, K.; Kirsch, S. A.; Han, J.; Sun, L.; Jiang, Z.; Unruh, T.; Böckmann, R. A. A Critical Comparison of Biomembrane Force Fields: Structure and Dynamics of Model DMPC, POPC, and POPE Bilayers. *J. Phys. Chem. B* **2016**, *120*, 3888−3903.
(66) Lewis, R. N. A. H.; Mak, N.; McElhaney, R. N. A differential scanning calorimetric study of the thermotropic phase behavior of model membranes composed of phosphatidylcholines containing linear saturated fatty acyl chains. *Biochemistry* **1987**, *26*, 6118−6126.
(67) Huang, C. H.; Lapides, J. R.; Levin, I. W. Phase-transition behavior of saturated, symmetric chain phospholipid bilayer dispersions determined by Raman spectroscopy: correlation between spectral and thermodynamic parameters. *J. Am. Chem. Soc.* **1982**, *104*, 5926−5930.
(68) Mabrey, S.; Sturtevant, J. M. Investigation of phase transitions of lipids and lipid mixtures by sensitivity differential scanning calorimetry. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 3862−3866.
(69) Davis, P. J.; Fleming, B. D.; Coolbear, K. P.; Keough, K. M. W. Gel to liquid-crystalline transition temperatures of water dispersions of two pairs of positional isomers of unsaturated mixed-acid phosphatidylcholines. *Biochemistry* **1981**, *20*, 3633−3636.
(70) Wang, Z. Q.; Lin, H. N.; Li, S.; Huang, C. H. Calorimetric studies and molecular mechanics simulations of monounsaturated phosphatidylethanolamine bilayers. *J. Biol. Chem.* **1994**, *269*, 23491−23499.
(71) Borle, F.; Seelig, J. Ca^{2+} binding to phosphatidylglycerol bilayers as studied by differential scanning calorimetry and ^{2}H- and ^{31}P-nuclear magnetic resonance. *Chem. Phys. Lipids* **1985**, *36*, 263−283.

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

(72) Berendsen, H. J.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(73) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*−*2*, 19−25.
(74) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(75) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(76) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952−962.
(77) Poger, D.; Caron, B.; Mark, A. E. Validating lipid force fields against experimental data: Progress, challenges and perspectives. *Biochim. Biophys. Acta, Biomembr.* **2016**, *1858*, 1556−1565.
(78) Braun, A. R.; Brandt, E. G.; Edholm, O.; Nagle, J. F.; Sachs, J. N. Determination of electron density profiles and area from simulations of undulating membranes. *Biophys. J.* **2011**, *100*, 2112−2120.
(79) Chiu, S.-W.; Pandit, S. A.; Scott, H. L.; Jakobsson, E. An Improved United Atom Force Field for Simulation of Mixed Lipid Bilayers. *J. Phys. Chem. B* **2009**, *113*, 2748−2763.
(80) Douliez, J.-P.; Ferrarini, A.; Dufourc, E.-J. On the relationship between C-C and C-D order parameters and its use for studying the conformation of lipid acyl chains in biomembranes. *J. Chem. Phys.* **1998**, *109*, 2513−2518.
(81) Piggot, T. J.; Allison, J. R.; Sessions, R. B.; Essex, J. W. On the Calculation of Acyl Chain Order Parameters from Lipid Simulations. *J. Chem. Theory Comput.* **2017**, *13*, 5683−5696.
(82) Ding, W.; Palaiokostas, M.; Wang, W.; Orsi, M. Effects of Lipid Composition on Bilayer Membranes Quantified by All-Atom Molecular Dynamics. *J. Phys. Chem. B* **2015**, *119*, 15263−15274.
(83) Gurtovenko, A. A.; Vattulainen, I. Calculation of the electrostatic potential of lipid bilayers from molecular dynamics simulations: Methodological issues. *J. Chem. Phys.* **2009**, *130*, No. 215107.
(84) Einstein, A.; Fürt, R. *Investigations on the Theory of Brownian Movement*; Dover Publications, 1956; p 119.
(85) Johner, N.; Harries, D.; Khelashvili, G. Implementation of a methodology for determining elastic properties of lipid assemblies from molecular dynamics simulations. *BMC Bioinf.* **2016**, *17*, No. 161.
(86) Biasini, M.; Mariani, V.; Haas, J.; Scheuber, S.; Schenk, A. D.; Schwede, T.; Philippsen, A. OpenStructure: a flexible software framework for computational structural biology. *Bioinformatics* **2010**, *26*, 2626−2628.
(87) Aurenhammer, F. Franz, Voronoi diagrams − a survey of a fundamental geometric data structure. *ACM Comput. Surv.* **1991**, *23*, 345−405.
(88) Uppulury, K.; Coppock, P. S.; Kindt, J. T. Molecular Simulation of the DPPE Lipid Bilayer Gel Phase: Coupling between Molecular Packing Order and Tail Tilt Angle. *J. Phys. Chem. B* **2015**, *119*, 8725.
(89) Reißer, S.; Poger, D.; Stroet, M.; Mark, A. E. Real Cost of Speed: The Effect of a Time-Saving Multiple-Time-Stepping Algorithm on the Accuracy of Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2017**, *13*, 2367−2372.
(90) Anézo, C.; de Vries, A. H.; Höltje, H.-D.; Tieleman, D. P.; Marrink, S.-J. Methodological Issues in Lipid Bilayer Simulations. *J. Phys. Chem. B* **2003**, *107*, 9424−9433.
(91) Chowdhary, J.; Harder, E.; Lopes, P. E. M.; Huang, L.; MacKerell, A. D.; Roux, B. A Polarizable Force Field of Dipalmitoylphosphatidylcholine Based on the Classical Drude Model for Molecular Dynamics Simulations of Lipids. *J. Phys. Chem. B* **2013**, *117*, 9142−9160.
(92) Warshaviak, D. T.; Muellner, M. J.; Chachisvilis, M. Effect of membrane tension on the electric field and dipole potential of lipid bilayer membrane. *Biochim. Biophys. Acta, Biomembr.* **2011**, *1808*, 2608−2617.
(93) Vermeer, L. S.; de Groot, B. L.; Réat, V.; Milon, A.; Czaplicki, J. Acyl chain order parameter profiles in phospholipid bilayers: computation from molecular dynamics simulations and comparison with 2H NMR experiments. *Eur. Biophys. J.* **2007**, *36*, 919−931.
(94) Douliez, J. P.; Léonard, A.; Dufourc, E. J. Restatement of order parameters in biomembranes: calculation of C-C bond order parameters from C-D quadrupolar splittings. *Biophys. J.* **1995**, *68*, 1727−1739.
(95) Petrache, H. I.; Dodd, S. W.; Brown, M. F. Area per lipid and acyl length distributions in fluid phosphatidylcholines determined by (2)H NMR spectroscopy. *Biophys. J.* **2000**, *79*, 3172−3192.
(96) Borle, F.; Seelig, J. Hydration of Escherichia coli lipids: Deuterium T1 relaxation time studies of phosphatidylglycerol, phosphatidylethanolamine and phosphatidylcholine. *Biochim. Biophys. Acta, Biomembr.* **1983**, *735*, 131−136.
(97) Ulrich, A. S.; Volke, F.; Watts, A. The dependence of phospholipid head-group mobility on hydration as studied by deuterium-NMR spin-lattice relaxation time measurements. *Chem. Phys. Lipids* **1990**, *55*, 61−66.
(98) Ulrich, A.; Watts, A. Molecular response of the lipid headgroup to bilayer hydration monitored by 2H-NMR. *Biophys. J.* **1994**, *66*, 1441−1449.
(99) Lairion, F.; Filler, R.; Disalvo, E. Reversed micelles as model systems to study interfacial properties of lipid bilayers. *Colloids Surf., B* **2002**, *25*, 369−371.
(100) Semchyschyn, D. J.; Macdonald, P. M. Conformational response of the phosphatidylcholine headgroup to bilayer surface charge: torsion angle constraints from dipolar and quadrupolar couplings in bicelles. *Magn. Reson. Chem.* **2004**, *42*, 89−104.
(101) Dickson, C. J.; Madej, B. D.; Skjevik, A. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **2014**, *10*, 865−879.
(102) Doux, J. P. F.; Hall, B. A.; Killian, J. A. How lipid headgroups sense the membrane environment: an application of ^{1}4N NMR. *Biophys. J.* **2012**, *103*, 1245−1253.
(103) Büldt, G.; Gally, H. U.; Seelig, A.; Seelig, J.; Zaccai, G. Neutron diffraction studies on selectively deuterated phospholipid bilayers. *Nature* **1978**, *271*, 182−184.
(104) Botan, A.; Favela-Rosales, F.; Fuchs, P. F. J.; Javanainen, M.; Kanduč, M.; Kulig, W.; Lamberg, A.; Loison, C.; Lyubartsev, A.; Miettinen, M. S.; Monticelli, L.; Määttä, J.; Ollila, O. H. S.; Retegan, M.; Róg, T.; Santuz, H.; Tynkkynen, J. Toward Atomistic Resolution Structure of Phosphatidylcholine Headgroup and Glycerol Backbone at Different Ambient Conditions. *J. Phys. Chem. B* **2015**, *119*, 15075−15088.
(105) Venable, R. M.; Ingólfsson, H. I.; Lerner, M. G.; Perrin, B. S.; Camley, B. A.; Marrink, S. J.; Brown, F. L. H.; Pastor, R. W. Lipid and Peptide Diffusion in Bilayers: The Saffman-Delbrück Model and Periodic Boundary Conditions. *J. Phys. Chem. B* **2017**, *121*, 3443−3457.
(106) Vögele, M.; Köfinger, J.; Hummer, G. Hydrodynamics of Diffusion in Lipid Membrane Simulations. *Phys. Rev. Lett.* **2018**, *120*, No. 268104.
(107) Lindblom, G.; Orädd, G. Lipid lateral diffusion and membrane heterogeneity. *Biochim. Biophys. Acta, Biomembr.* **2009**, *1788*, 234−244.
(108) Poger, D.; Mark, A. E. Lipid Bilayers: The Effect of Force Field on Ordering and Dynamics. *J. Chem. Theory Comput.* **2012**, *8*, 4807−4817.
(109) Piggot, T. J.; Piñeiro, Á.; Khalid, S. Molecular Dynamics Simulations of Phosphatidylcholine Membranes: A Comparative Force Field Study. *J. Chem. Theory Comput.* **2012**, *8*, 4593−4609.
(110) Nagle, J. F. Experimentally determined tilt and bending moduli of single-component lipid bilayers. *Chem. Phys. Lipids* **2017**, *205*, 18−24.
(111) Doktorova, M.; Harries, D.; Khelashvili, G. Determination of bending rigidity and tilt modulus of lipid membranes from real-space

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

fluctuation analysis of molecular dynamics simulations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 16806−16818.
(112) Nagle, J. F.; Cognet, P.; Dupuy, F. G.; Tristram-Nagle, S. Structure of gel phase DPPC determined by X-ray diffraction. *Chem. Phys. Lipids* **2019**, *218*, 168−177.
(113) Sun, W.-J.; Suter, R. M.; Knewtson, M. A.; Worthington, C. R.; Tristram-Nagle, S.; Zhang, R.; Nagle, J. F. Order and disorder in fully hydrated unoriented bilayers of gel-phase dipalmitoylphosphatidylcholine. *Phys. Rev. E* **1994**, *49*, 4665−4676.
(114) McIntosh, T. J.; Simon, S. A. Area per molecule and distribution of water in fully hydrated dilauroylphosphatidylethanolamine bilayers. *Biochemistry* **1986**, *25*, 4948−4952.
(115) Marinov, R.; Dufourc, E. J. Thermotropism and hydration properties of POPE and POPE-cholesterol systems as revealed by solid state2H and31P-NMR. *Eur. Biophys. J.* **1996**, *24*, 423−431.
(116) Rand, R. P.; Fuller, N.; Parsegian, V. A.; Rau, D. C. Variation in hydration forces between neutral phospholipid bilayers: evidence for hydration attraction. *Biochemistry* **1988**, *27*, 7711−7722.
(117) Hwang, P. M.; Zhou, N.; Shan, X.; Arrowsmith, C. H.; Vogel, H. J. Three-Dimensional Solution Structure of Lactoferricin B, an Antimicrobial Peptide Derived from Bovine Lactoferrin †. *Biochemistry* **1998**, *37*, 4288−4298.
(118) Nguyen, L. T.; Schibli, D. J.; Vogel, H. J. Structural studies and model membrane interactions of two peptides derived from bovine lactoferricin. *J. Pept. Sci. A.* **2005**, *11*, 379−389.
(119) Schibli, D. J.; Hwang, P. M.; Vogel, H. J. The structure of the antimicrobial active center of lactoferricin B bound to sodium dodecyl sulfate micelles. *FEBS Lett.* **1999**, *446*, 213−217.
(120) Sohlenkamp, C.; Geiger, O. Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol. Rev.* **2016**, *40*, 133−159.
(121) Ma, H.; Cummins, D. D.; Edelstein, N. B.; Gomez, J.; Khan, A.; Llewellyn, M. D.; Picudella, T.; Willsey, S. R.; Nangia, S. Modeling Diversity in Structures of Bacterial Outer Membrane Lipids. *J. Chem. Theory Comput.* **2017**, *13*, 811−824.
(122) Castelletto, V.; de Santis, E.; Alkassem, H.; Lamarre, B.; Noble, J. E.; Ray, S.; Bella, A.; Burns, J. R.; Hoogenboom, B. W.; Ryadnov, M. G. Structurally plastic peptide capsules for synthetic antimicrobial viruses. *Chem. Sci.* **2016**, *7*, 1707−1711.
(123) Kandasamy, S. K.; Larson, R. G. Binding and insertion of α-helical anti-microbial peptides in POPC bilayers studied by molecular dynamics simulations. *Chem. Phys. Lipids* **2004**, *132*, 113−132.
(124) Vivcharuk, V.; Tomberli, B.; Tolokh, I. S.; Gray, C. G. Prediction of binding free energy for adsorption of antimicrobial peptide lactoferricin B on a POPC membrane. *Phys. Rev. E* **2008**, *77*, No. 031913.
(125) Wang, Y.; Schlamadinger, D. E.; Kim, J. E.; McCammon, J. A. Comparative molecular dynamics simulations of the antimicrobial peptide CM15 in model lipid bilayers. *Biochim. Biophys. Acta, Biomembr.* **2012**, *1818*, 1402−1409.
(126) Strøm, M. B.; Haug, B. E.; Rekdal, O.; Skar, M. L.; Stensen, W.; Svendsen, J. S. Important structural features of 15-residue lactoferricin derivatives and methods for improvement of antimicrobial activity. *Biochem. Cell Biol.* **2002**, *80*, 65−74.
(127) Chan, D. I.; Prenner, E. J.; Vogel, H. J. Tryptophan- and arginine-rich antimicrobial peptides: Structures and mechanisms of action. *Biochim. Biophys. Acta, Biomembr.* **2006**, *1758*, 1184−1202.
(128) Petrov, D.; Margreitter, C.; Grandits, M.; Oostenbrink, C.; Zagrovic, B. A Systematic Framework for Molecular Dynamics Simulations of Protein Post-Translational Modifications. *PLoS Comput. Biol.* **2013**, *9*, No. e1003154.

DOI: 10.1021/acs.jctc.9b00509
J. Chem. Theory Comput. 2019, 15, 5175−5193

Supporting Information:

Lipid head group parameterisation for GROMOS 54a8: a consistent approach with protein force field description

Irene Marzuoli, Christian Margreitter and Franca Fraternali

Randall Centre for Cell and Molecular Biology, King's College London, London

List of Tables

List of Figures

1 Supplementary results: dipole potential

For the five phospholipids and the five different parameters tested, we computed the dipole potential across the membrane, according to the procedure outlined in Section 2.5 of the paper.

The results are outlined in Figure 1 and show a striking difference between the profiles obtained with the Chiu partial charges set and the Reif-Margreitter one: in particular, while the former suggests local maxima of the dipole potential in the head regions, the latter results in a potential rising to its higher value in the center of the tails region.

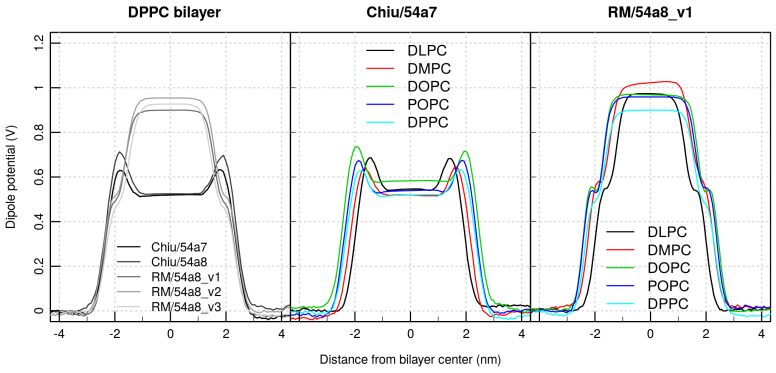


Figure 1: Dipole potential for respectively: (left) a DPPC bilayer, simulated with the five parameter sets explored in the manuscript; (center) the five different phospholipid considered and parameter set Chiu/54a7; (right) the five different phospholipid considered and parameter set RM/54a8_v1.

The results for the Chiu set of partial charges (and either the 54a7 or 54a8 GROMOS force field) are compatible with what obtained previously by Gurtovenko et al. in Ref. [7] using the same charges and the Berger force field (a slight modification of GROMOS tailored for lipids simulations). On the contrary, the new results are qualitatively more similar to what obtained with the all atom force field CHARMM36 in Ref. [4] and [23], as they both suggest a peak of the potential in the lipid tail regions. However, with respect to the CHARMM36 profiles, the ones obtained by the RM/54a8_v1 parameters give a flatter trend in the center

of the bilayer. Moreover, the peaks relative to the phosphate regions are shallower for the RM/54a8_v1 simulations with respect to the CHARMM ones.

To discriminate between the two behaviours it is necessary a comparison with experimental results. Few works compute the profile of the dipole potential along the bilayer normal, focussing instead on measuring the difference between the dipole potential on the two sides of the bilayer [20, 19, 13]. Never the less, Wang et al. [22] measured its profile for ester-DPhPC lipid vesicles, suggesting a peak in the central region of the bilayer. This lipid is derived from DPPC by adding an ester linkage: as such, its profile will not be identical to the one produced by DPPC, but being the modification in the ester region only, we can assume that the qualitative trend remains the same and thus can be used for the present comparison.

Additionally to experimental data, theoretical profiles obtained from ab initio calculations and geometrical assumptions on the structure of lipids bilayers suggest indeed a plateau in the tail region [3] - or at the very least do not support the presence of peaks in the head regions, as showed when the Chiu charge set is used in the simulations.

For all the above reasons we once more consider the new set of parameter (RM/54a8_v1) as a better compromise than the previous sets Chiu/54a7 or Chiu/54a8.

Supplementary Movie 1: file LFC_54a8_v1_0.mp4

The movie displays the simulation of LFC on a bacterial membrane (composition DLPC:DLPG with 3:1 ratio) for both parameter set Chiu/54a8 (left half) and RM/54a8_v1 (right half), starting from orientation OII. The lipid choline and phosphate atoms are shown as van der Waals beads, the protein as cyan cartoon and bonds coloured by residue type (blue positive, green polar, white hydrophobic) except for the two Tryptophan residues which are coloured in orange.

Control simulations

Sim	Charges [a]	FF	Long range ES [b]	Nr lipids	T (K)
0	RM	54A7	PME	512	323
RF	Chiu	54A7	RF	512	323
Small	Chiu	54A7	PME	128	323
303	RM	54A8_v1	PME	512	303
333	RM	54A8_v1	PME	512	333

Table 1: Table of additional control simulations, run on the DPPC bilayer. [a] Charge set: Chiu from Ref [2], RM Reif-Margreitter as illustrated in the present work. [b] RF reaction field [21], PME Particle Mesh Ewald summation [6].

Simulations of POPE

Sim	Charges [a]	FF	Ester carbon type
1CH0	PC	54a8	CH0
CH0	RM	54a8	CH0

Table 2: Table of simulations for POPE. All are run for 500 ns and systems consisting of 512 lipid molecules (256 per layer), using a PME long range electrostatic scheme. [a] Charge set: PC Piggot-Chiu from Ref [14], RM Reif-Margreitter as illustrated in the present work.

K_A (mN m^{-1})

ID	Charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	553	477	607	515	499
2	Chiu/54a8	471	500	471	679	612
3	RM/54a8_v1	562	354	485	397	329
4	RM/54a8_v2	365	318	518	466	406
5	RM/54a8_v3	332	300	426	345	376
Experiment [a]		234(23)	188-265(18)	231(20)	180-330	248(20)

Table 3: Average isothermal area compressibility module K_A over the last 300 ns of simulations for phosphocholine bilayers. All simulations run at 303 K, except for DPPC (323 K). [a] Values from Ref. [9] and Table 1 in Ref. [15], in parenthesis the standard error on the last digit is shown.

Hydrophobic thickness D_{HH} (nm)

ID	Charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	2.83	3.59	3.05	3.30	4.30
2	Chiu/54a8	2.78	3.52	2.90	3.14	4.21
3	RM/54a8_v1	2.72	3.48	2.89	3.22	4.06
4	RM/54a8_v2	2.61	3.33	2.82	3.08	4.06
5	RM/54a8_v3	2.38	3.13	2.65	2.86	3.77
Experiment [a]		3.08	3.44-3.60	3.53-3.71	3.70	3.42-3.83

Luzzati thickness D_B (nm)

ID	Charges/FF	DLPC	DMPC	DOPC	POPC	DPPC
1	Chiu/54A7	3.11	3.54	4.13	4.00	3.93
2	Chiu/54a8	3.14	3.61	3.97	3.96	3.96
3	RM/54a8_v1	3.04	3.48	3.94	3.88	3.75
4	RM/54a8_v2	3.00	3.34	3.87	3.77	3.74
5	RM/54a8_v3	2.80	3.21	3.62	3.57	3.50
Experiment [a]		3.14	3.63-3.96	3.59-3.87	3.68	3.50-3.83

Table 4: Bilayer thickness, according to the Phosphate or Luzzati methods, derived from the electron density profiles of phosphocholine bilayers. All simulations run at 303 K, except for DPPC (323 K). [a] Values from Ref. [16] and Table 1 in Ref. [15]

Nr water molecules per group

Group	ID	DLPC	DMPC	DOPC	DPPC	POPC
All	1	3.9 (13.2)	3.8 (12.9)	3.9 (13.3)	3.7 (12.7)	3.8 (13.1)
	2	3.6 (12.7)	3.5 (12.4)	3.6 (12.9)	3.4 (12.3)	3.6 (12.6)
	3	4.3 (14.1)	4.3 (13.8)	4.4 (14.3)	4.2 (13.7)	4.3 (14.1)
	4	4.6 (15.1)	4.6 (15.0)	4.6 (15.1)	4.5 (14.7)	4.6 (15.1)
	5	5.1 (16.9)	5.1 (16.8)	5.0 (16.9)	4.9 (16.6)	5.1 (16.8)
Cho	1	13.1	12.8	13.3	12.6	13.1
	2	12.6	12.3	12.7	12.2	12.5
	3	14.3	14.1	14.5	13.9	14.3
	4	15.3	15.2	15.4	14.9	15.3
	5	17.0	16.9	17.0	16.7	17.0
PO4	1	3.1 (9.8)	3.0 (9.6)	3.1 (9.9)	3.1 (9.6)	3.1 (9.8)
	2	2.6 (9.3)	2.5 (9.0)	2.6 (9.4)	2.6 (9.2)	2.5 (9.2)
	3	4.0 (9.8)	4.0 (9.6)	4.0 (9.9)	3.9 (9.8)	4.0 (9.8)
	4	4.3 (10.3)	4.3 (10.3)	4.3 (10.4)	4.2 (10.3)	4.3 (10.4)
	5	4.8 (11.5)	4.7 (11.5)	4.8 (11.5)	4.7 (11.5)	4.8 (11.5)
Gly	1	0.3 (3.5)	0.3 (3.4)	0.3 (3.5)	0.3 (3.4)	0.3 (3.5)
	2	0.3 (3.4)	0.2 (3.2)	0.3 (3.4)	0.2 (3.3)	0.2 (3.3)
	3	0 (3.5)	0 (3.5)	0 (3.6)	0 (3.4)	0 (3.5)
	4	0 (3.8)	0 (3.8)	0 (3.8)	0 (3.7)	0 (3.8)
	5	0 (4.3)	0 (4.3)	0 (4.2)	0 (4.2)	0 (4.3)
CO1	1	1.4	1.4	1.4	1.4	1.4
	2	1.5	1.5	1.5	1.5	1.5
	3	1.2	1.2	1.2	1.1	1.2
	4	1.3	1.2	1.3	1.1	1.3
	5	1.3	1.3	1.3	1.3	1.3
CO2	1	0.5	0.5	0.5	0.5	0.5
	2	0.6	0.6	0.6	0.6	0.6
	3	0.7	0.6	0.7	0.6	0.7
	4	0.7	0.7	0.7	0.7	0.7
	5	0.8	0.8	0.8	0.8	0.8

Table 5: Integrated value of water molecules around each lipid and for their constituting chemical groups (Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions). The integration is performed up to the first peak and, if applicable, to the second (the latter of which is given in brackets). Peak positions are the same throughout the simulations, except for glycerol in ID3 to ID5, where the first peak is not shown (thus the value of 0, see main text for interpretation).

POPE (313 K) structural properties

ID	Charges/FF/ester	ApL (nm^2)	K_A ($mN\ m^{-1}$)	D_{HH} (nm)	D_B (nm)
1CH0	Chiu/54a8/CH0	0.591(5)	437	2.48	3.92
CH0	RM/54a8/CH0	0.595(5)	361	2.52	3.89
Experiment [a]		0.566	–	–	4.13

Table 6: Area per lipid, lateral compressibility and bilayer thickness, according to the phosphate or Luzzati methods, derived from the electron density profiles for POPE bilayer. [a] From [17].

POPE (313 K) Nr water molecules per group

ID	Charges/FF/ester	All	Am	PO4	Gly	CO1	CO2
1CH0	Chiu/54a8/CH0	1.6 (5.7)	2.2	3.3 (10.6)	0 (3.4)	0.8	0.5
CH0	RM/54a8/CH0	2.1 (6.4)	2.5	4.2 (11.3)	0 (3.6)	0.6	0.6

Table 7: Number of water molecules per lipid hydrating the POPE bilayer and its different polar moieties (Am: amine, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions). Integration was performed up to the first peak and when relevant to the second one (value in bracket). Peak positions are the same for both simulations.

POPG (303 K) structural properties

ID	Charges/FF/ester	ApL (nm^2)	K_A ($mN\ m^{-1}$)	D_{HH} (nm)	D_B (nm)
CH0	RM/54a8/CH0	0.646(7)	198	3.76	3.98
Experiment [a]		0.661(13)	–	3.74(7)	–

Table 8: Area per lipid, lateral compressibility and bilayer thickness, according to the phosphate or Luzzati methods, derived from the electron density profiles. [a] From [11].

POPG (303 K) Nr water molecules per group

ID	Charges/FF/ester	All	Etam	PO4	Gly	CO1	CO2
H0	RM/54a8/CH0	0.2 (3.7)	0.5 (7.8)	3.9	3.2 (6.2)	0.5	0.4

Table 9: Number of water molecules per lipid hydrating the POPG bilayer and its different polar moieties (Etam: etamine, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions). Integration was performed up to the first peak, and when relevant to the second one (value in bracket). Peak positions are the same for both simulations.

Comparison with the previous parameterization work

	Poger et al. [16, 15]		Present work	
Lipid parameters				
Choline partial charges	Chiu	[2]	Reif	[18]
Phosphate partial charges	Chiu	[2]	Margreitter	[8]
CH3p-OM C12 [$kJ\,mol^{-1}nm^{12}$]	$1.58E$-5	[16]	$1.10E$-5	Present work
CH3p-CH3p C12 [$kJ\,mol^{-1}nm^{12}$]	$2.66E$-5	[10]	$6.48E$-5	[18]
Simulations details				
GROMACS version	3.2.1		2016.3	[1]
Long range electrostatics	Reaction Field	[21]	PME	[6]
Cut-off [a]	Twin-range		Single	
Short range cut-off [a]	0.8 nm		NA	
Long range cut-off	1.4 nm		1.4 nm	
System size (# of lipids)	128		512	
Sampling time	120 ns		300 ns	
# replicas	2		1	

Table 10: Table of the changes occurred between simulations run in Poger et al. [16, 15] and the present work, in terms of parameters used and simulation conditions. [a] Settings not reproducible with GROMACS version 2016.3 used in the present work.

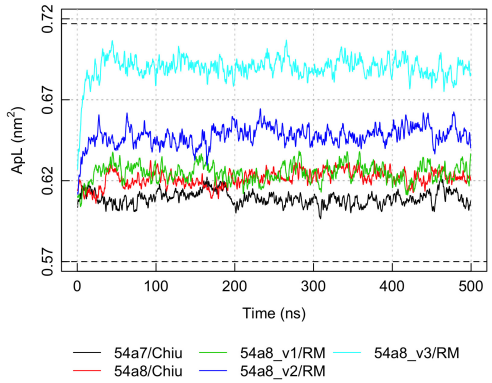


Figure 2: Time series of the area per lipid for a DPPC bilayer, for simulation ID 1 to 5 (see Table 1 for details).

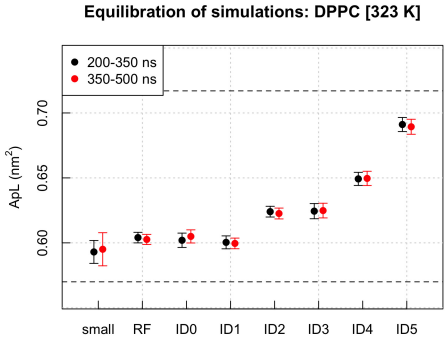


Figure 3: Area per lipid (computed from the simulation box sides) of a DPPC bilayer, for the five test simulation (ID 1 to 5 in Table 1) and the control ones (SI Table 1), computed for two non overlapping time intervals.

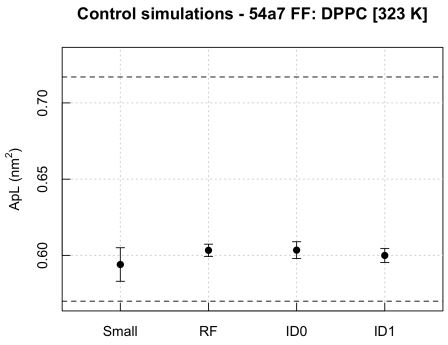


Figure 4: Area per lipid (computed from the simulation box sides) for a DPPC bilayer, for the control simulations as described in SI Table 1). Simulation ID1 (see Table 1) has been included for comparison, as each control simulation (small, RF, ID0) differs from ID1 by only one property (either the system size, long-range electrostatic treatment or the head group partial charges).

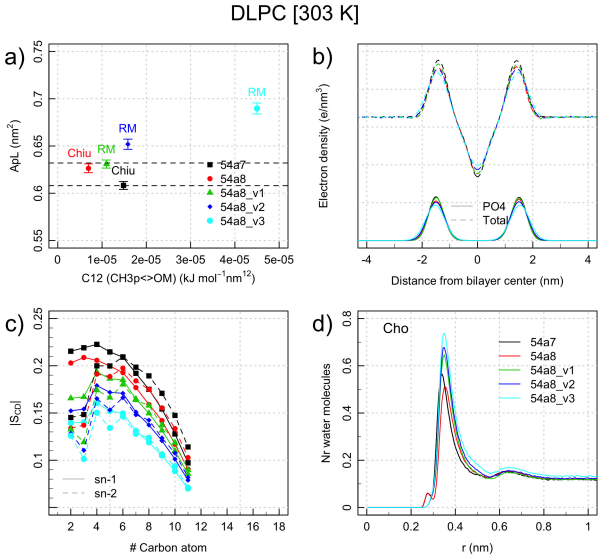


Figure 5: Area per lipid (a), electron density of CH2 tails and total (b), order parameter (c) and distribution of the distance between water oxygen and the nearest lipid Choline atom for a DLPC bilayer (d), obtained by 5 different sets of parameters (see legend in panel (a,c) and Table 1).

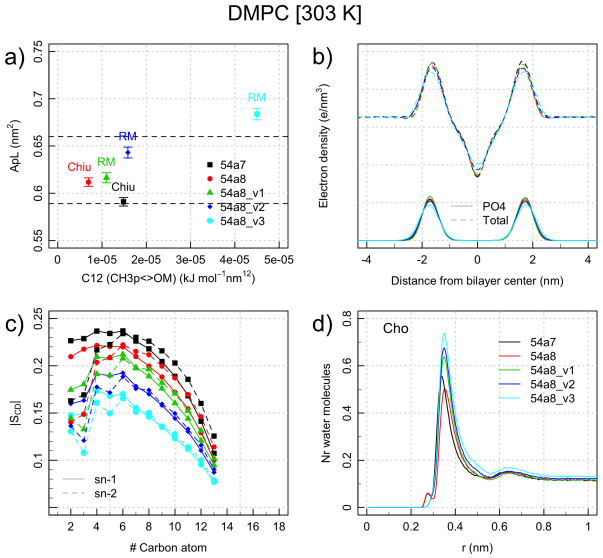


Figure 6: Area per lipid (a), electron density of CH2 tails and total (b), order parameter (c) and distribution of the distance between water oxygen and the nearest lipid Choline group atom for a DMPC bilayer (d), obtained by 5 different sets of parameters (see legend in panel (a,c) and Table 1).

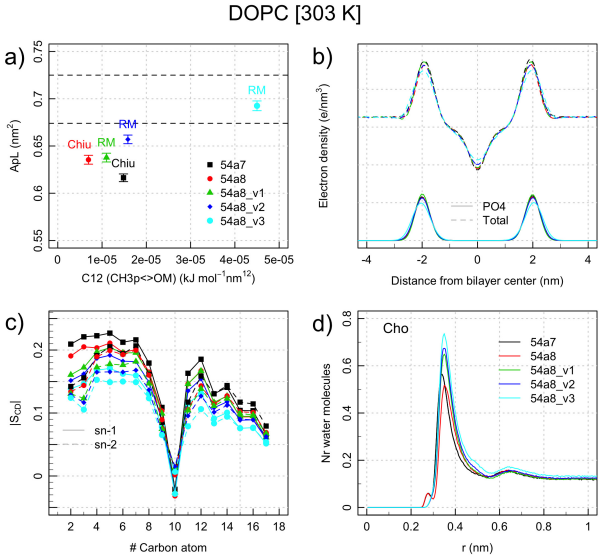


Figure 7: Area per lipid (a), electron density of CH2 tails and total (b), order parameter (c) and distribution of the distance between water oxygen and the nearest lipid Choline group atom for a DOPC bilayer (d), obtained by 5 different sets of parameters (see legend in panel (a,c) and Table 1).

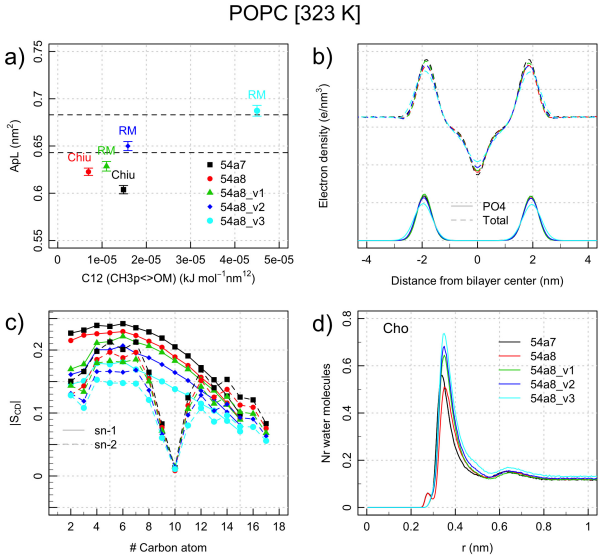


Figure 8: Area per lipid (a), electron density of CH2 tails and total (b), order parameter (c) and distribution of the distance between water oxygen and the nearest lipid Choline group atom for a POPC bilayer (d), obtained by 5 different sets of parameters (see legend in panel (a,c) and Table 1).

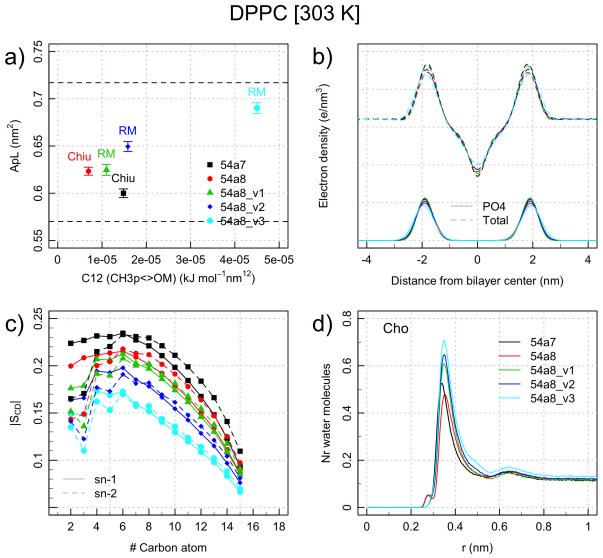


Figure 9: Area per lipid (a), electron density of CH2 tails and total (b), order parameter (c) and distribution of the distance between water oxygen and the nearest lipid Choline group atom for a DPPC bilayer (d), obtained by 5 different sets of parameters (see legend in panel (a,c) and Table 1).

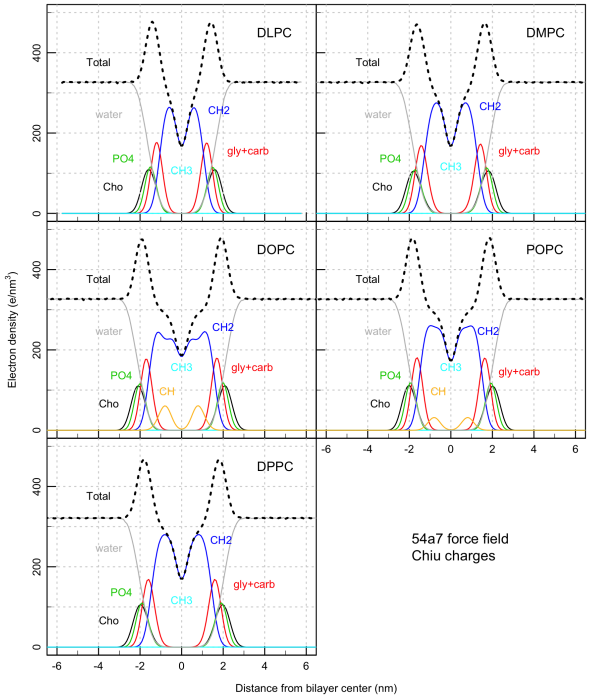


Figure 10: Electron density profiles of the whole hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers (Total) and of their individual components (Cho: choline, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of the acyl chains, CH: CHdCH groups in the oleoyl chains, CH3: terminal methyls of the acyl chains) for simulation ID1 (54A7 force field, Chiu charges).

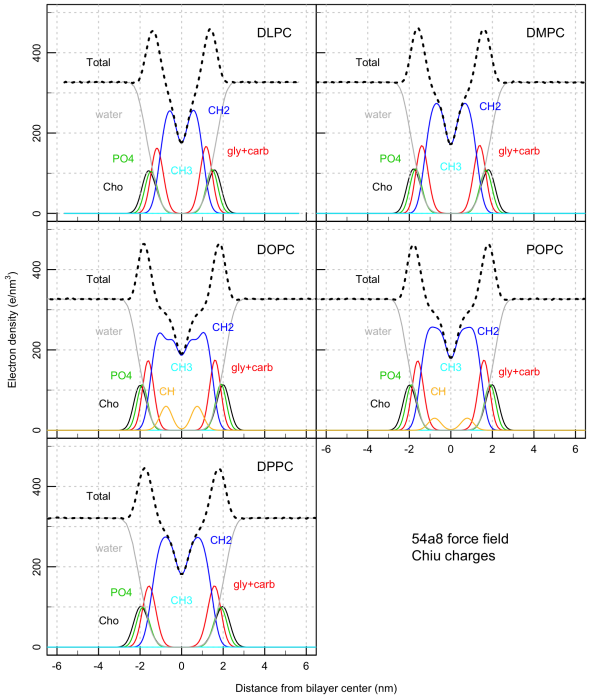


Figure 11: Electron density profiles of the whole hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers (Total) and of their individual components (Cho: choline, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of the acyl chains, CH: CHdCH groups in the oleoyl chains, CH3: terminal methyls of the acyl chains) for simulation ID2 (54A8 force field, Chiu charges).

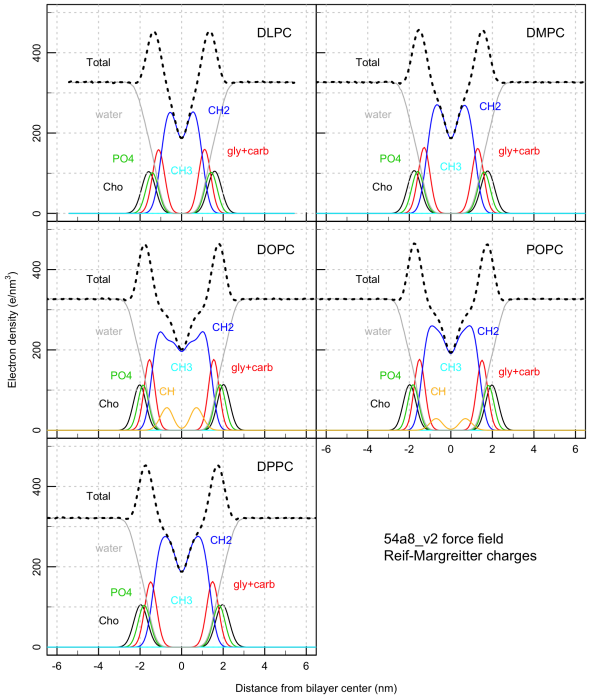


Figure 12: Electron density profiles of the whole hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers (Total) and of their individual components (Cho: choline, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of the acyl chains, CH: CHdCH groups in the oleoyl chains, CH3: terminal methyls of the acyl chains) for simulation ID4 (54A8_v2 force field, Reif-Margreitter charges).

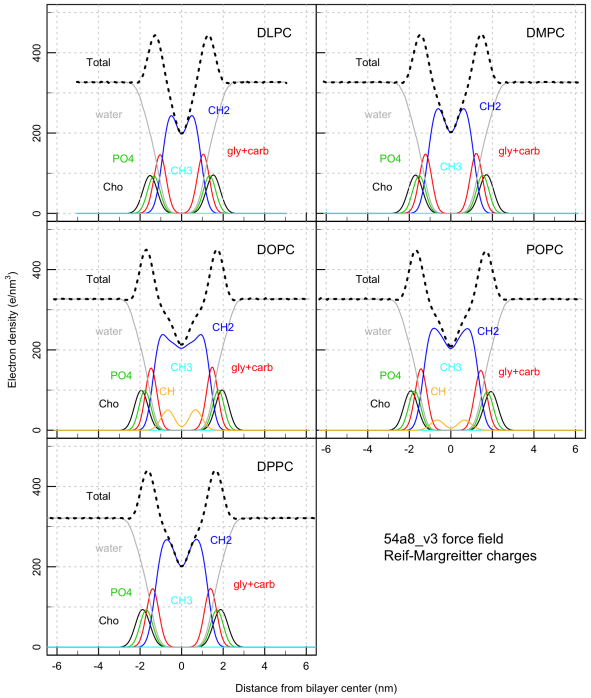


Figure 13: Electron density profiles of the whole hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers (Total) and of their individual components (Cho: choline, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of the acyl chains, CH: CHdCH groups in the oleoyl chains, CH3: terminal methyls of the acyl chains) for simulation ID5 (54A8_v3 force field, Reif-Margreitter charges).

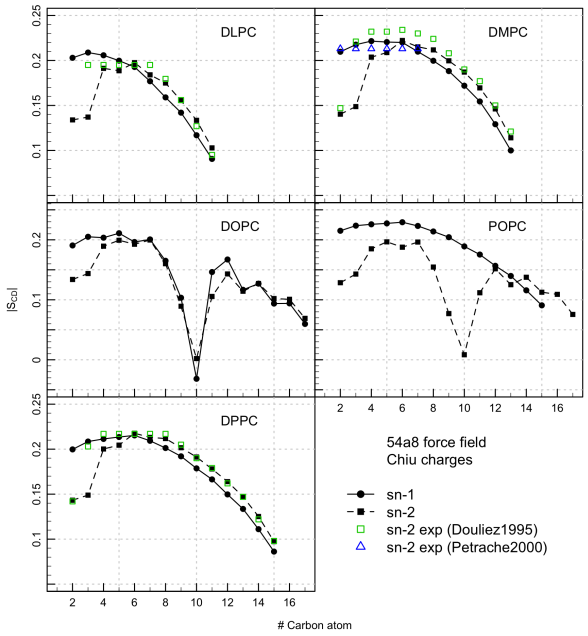


Figure 14: Deuterium order parameter S_{CD} profiles of the sn-1 and sn-2 fatty acyl chains of hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers calculated from simulations ID2 (54A8 force field, Chiu charges). The S_{CD} values are averaged over all the lipid sn-1 and -2 acyl chains in the systems (proS hydrogen only). Experimental values: Douliez1995 from Ref. [5], Petrache2000 from Ref. [12].

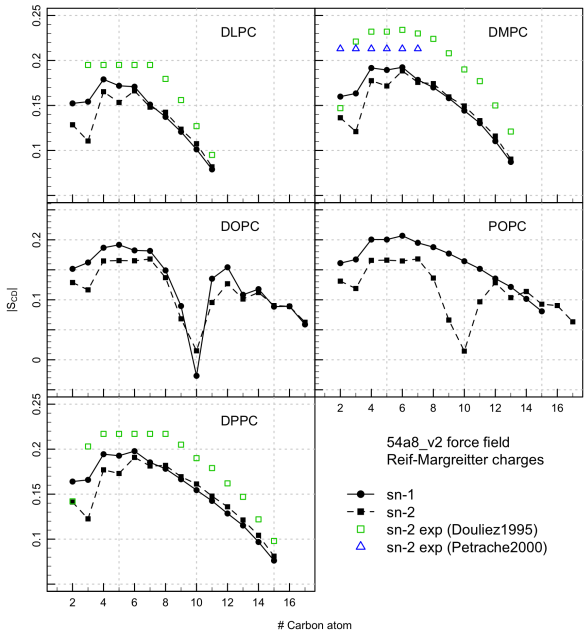


Figure 15: Deuterium order parameter S_{CD} profiles of the sn-1 and sn-2 fatty acyl chains of hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers calculated from simulations ID4 (54A8_v2 force field, Reif-Margreitter charges). The S_{CD} values are averaged over all the lipid sn-1 and -2 acyl chains in the systems (proS hydrogen only). Experimental values: Douliez1995 from Ref. [5], Petrache2000 from Ref. [12].

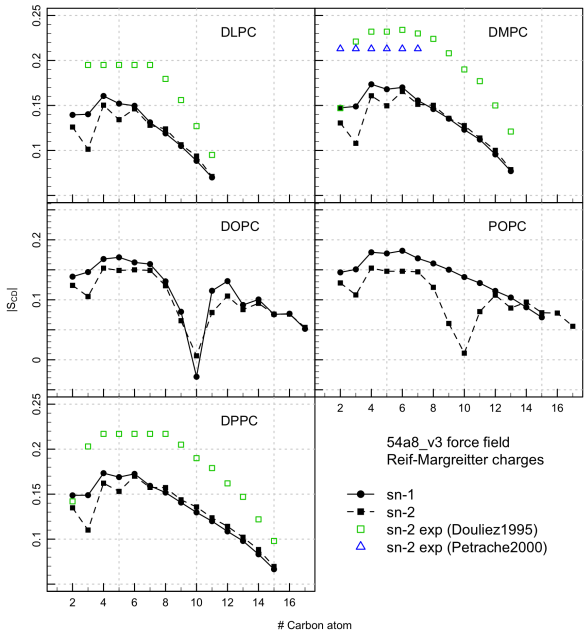


Figure 16: Deuterium order parameter S_{CD} profiles of the sn-1 and sn-2 fatty acyl chains of hydrated DLPC, DMPC, DOPC, DPPC and POPC bilayers calculated from simulations ID5 (54A8_v3 force field, Reif-Margreitter charges). The S_{CD} values are averaged over all the lipid sn-1 and -2 acyl chains in the systems (proS hydrogen only). Experimental values: Douliez1995 from Ref. [5], Petrache2000 from Ref. [12].

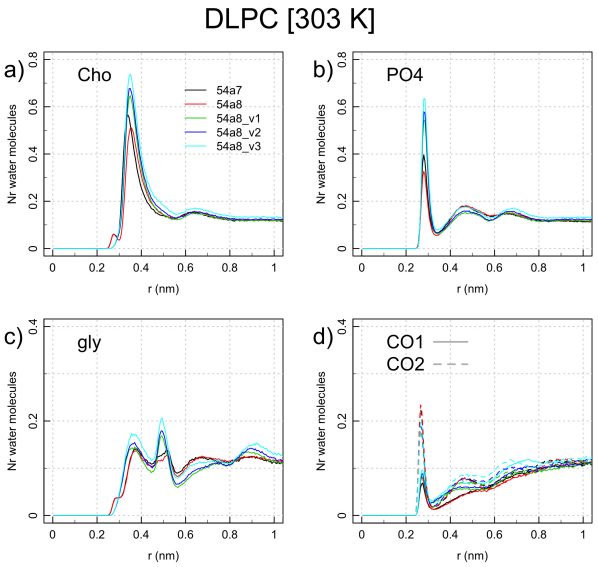


Figure 17: Distribution of the distance between the water oxygen and the nearest lipid headgroup atom in given moieties for simulations of a DLPC bilayer. Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions.

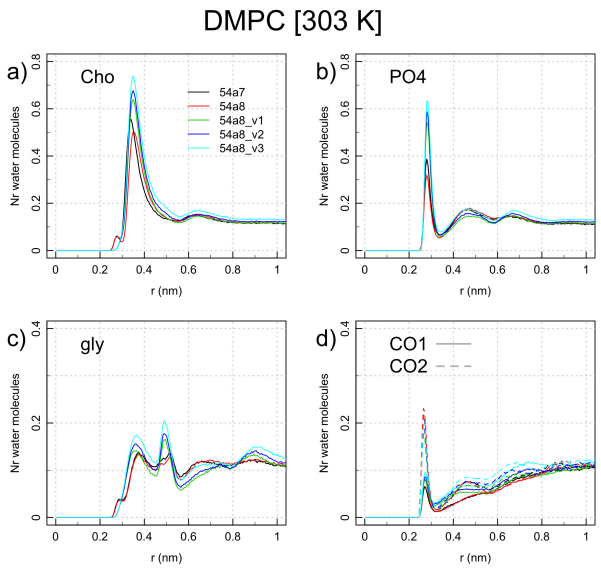


Figure 18: Distribution of the distance between the water oxygen and the nearest lipid headgroup atom in given moieties for simulations of a DMPC bilayer. Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions.

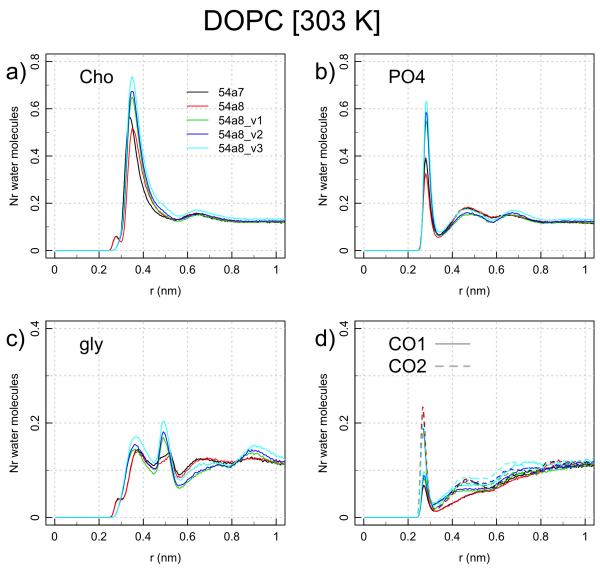


Figure 19: Distribution of the distance between the water oxygen and the nearest lipid headgroup atom in given moieties for simulations of a DOPC bilayer. Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions.

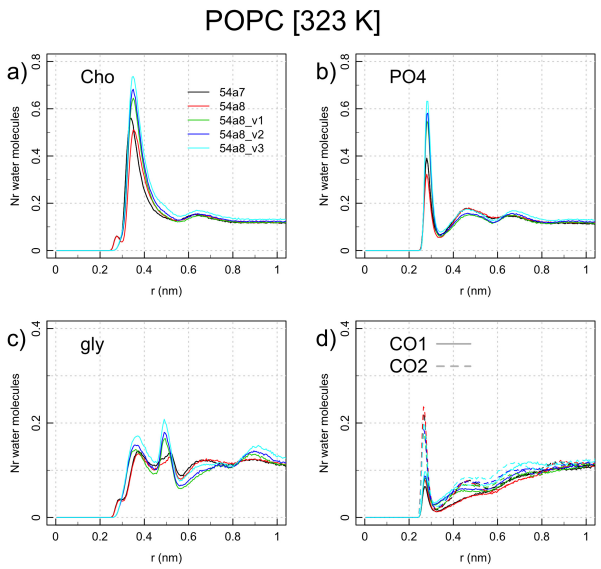


Figure 20: Distribution of the distance between the water oxygen and the nearest lipid headgroup atom in given moieties for simulations of a POPC bilayer. Cho: choline, PO4: phosphate, Gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions.

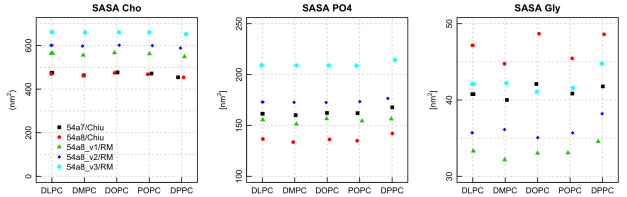


Figure 21: Values of SASA for Choline, Phosphate and Glycerol moieties for all the five phosphocholine and each of the parameter sets simulated.

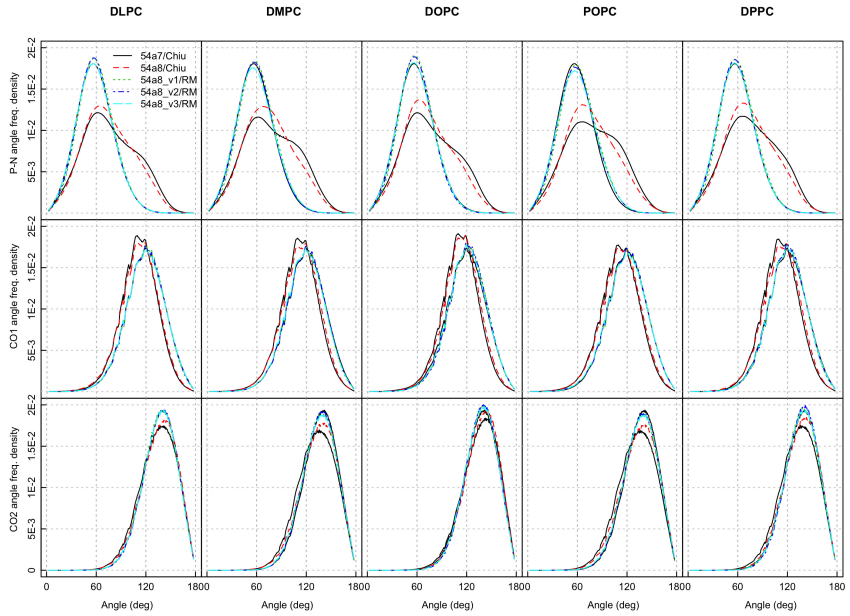


Figure 23: Distribution of the P-N, CO1 and CO2 angles with respect to the outward bilayer normal across different simulations for each lipid.

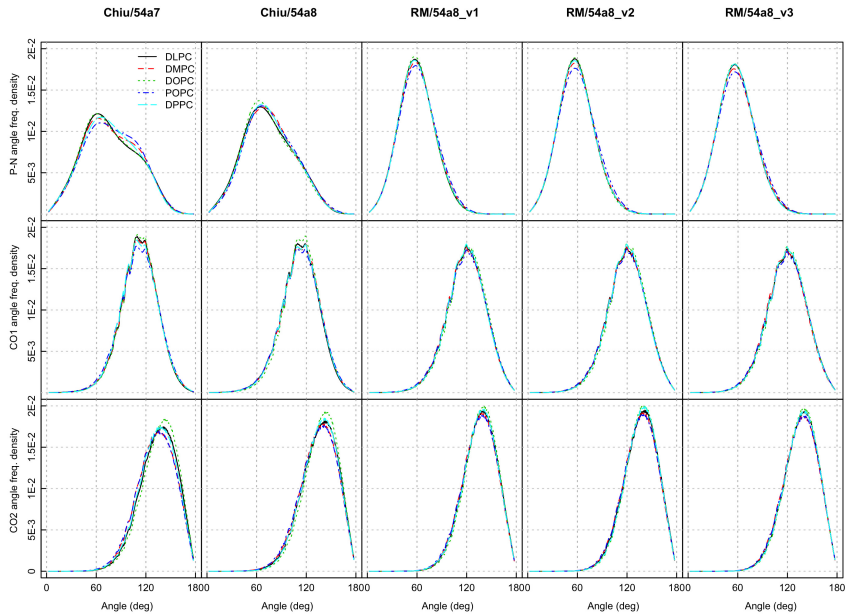


Figure 22: Distribution of the P-N, CO1 and CO2 angles with respect to the outward bilayer normal across different lipids for simulations ID 1 to 5.

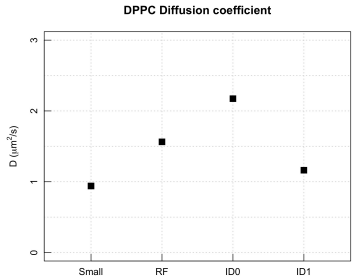


Figure 24: Lateral diffusion coefficient for DPPC in the control simulations as per SI Table 1.

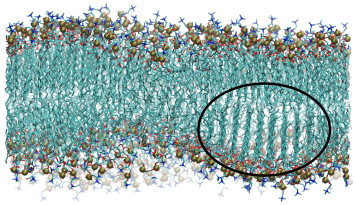


Figure 25: Snapshot after 400 ns simulation of a DPPC patch at 303 K with the 54a8_v1 parameter set. In the circle a detail of the ordered lipid tails which have already transitioned in a gel phase.

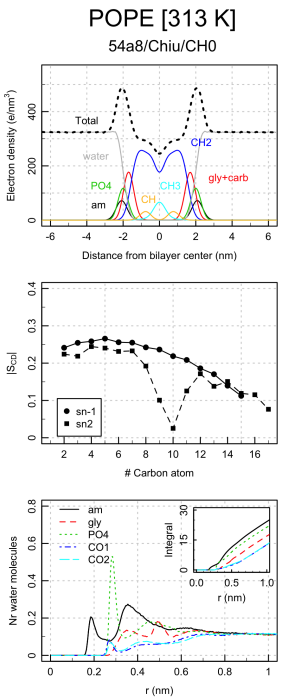


Figure 26: Properties of POPE bilayer simulated with the 54A8 force field, Chiu charges and CH0 esters. Top: electron density profiles (Total: whole bilayer, am: amine, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of acyl chains, CH: CHdCH groups of oleoyl chains, CH3: terminal methyls); middle: deuterium order parameter S_{CD} of the sn-1 and sn-2 fatty acyl chains (proS hydrogen only); bottom: distribution of the distance between the water oxygen and the nearest lipid headgroup atom - insets: integral of the distribution (am: amine, PO4: phosphate, gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions).

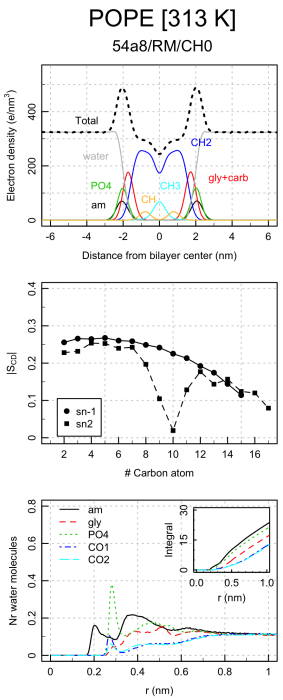


Figure 27: Properties of POPE bilayer simulated with the 54A8 force field, RM charges and CH0 esters. Top: electron density profiles (Total whole bilayer, am amide, PO4 phosphate, gly+carb glycerol carbonyl groups, CH2 methylenes of acyl chains, CH CHdCH groups of oleoyl chains, CH3 terminal methyls); middle: deuterium order parameter S_{CD} of the sn-1 and sn-2 fatty acyl chains (proS hydrogen only); bottom: distribution of the distance between the water oxygen and the nearest lipid headgroup atom - insets: integral of the distribution (am: amine, PO4: phosphate, gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions).

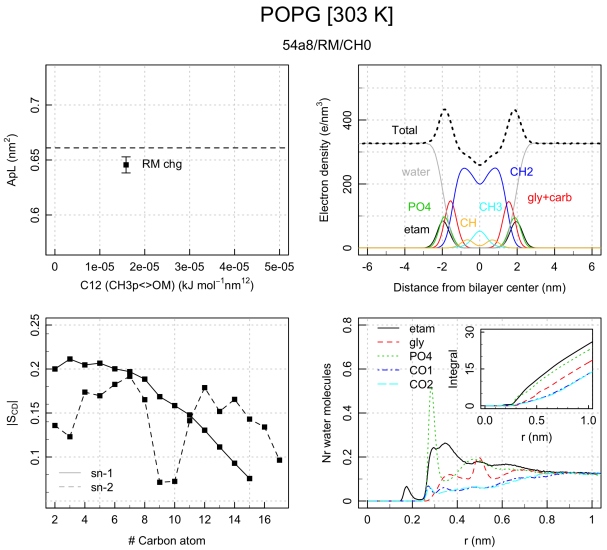


Figure 28: Properties of POPG bilayer simulated with the 54A8 force field, RM charges and CH0 esters. Top left: area per lipid and experimental value (dashed line); top right: electron density profiles (Total: whole bilayer, etam: etamine, PO4: phosphate, gly+carb: glycerol carbonyl groups, CH2: methylenes of acyl chains, CH: CHdCH groups of oleoyl chains, CH3: terminal methyls); bottom left: deuterium order parameter S_{CD} of the sn-1 and sn-2 fatty acyl chains (proS hydrogen only); bottom right: distribution of the distance between the water oxygen and the nearest lipid headgroup atom - insets: integral of the distribution (etam: etamine, PO4: phosphate, gly: glycerol, CO1 and CO2: carbonyl groups at the sn-1 and sn-2 positions).

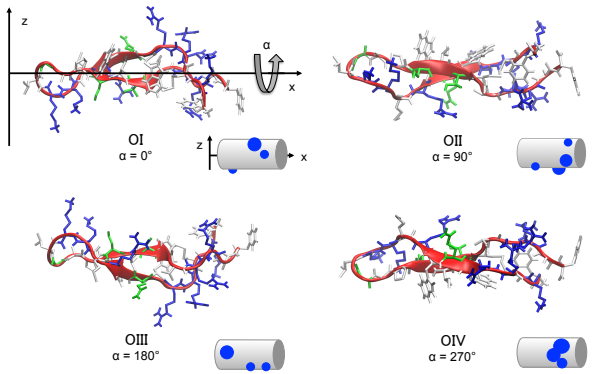


Figure 29: Four different orientations of the peptide 1LFC (bovine lactoferricin) used to sample different initial conditions for the peptide-membrane simulations. OII, OIII and OIV are obtained from OI with an anticlockwise rotation along the main axis (shown as x-axis in OI) of respectively 90°, 180°, 270°. The membrane plane is placed parallel to the main axis of 1LFC; the residues displayed in the bottom are the ones facing the lipids. Backbone red cartoon representation and whole peptide bonds representation coloured by residue type. The insets shows a cartoon representation of the portion of LFC facing the membrane, with blue dots showing the positive patches.

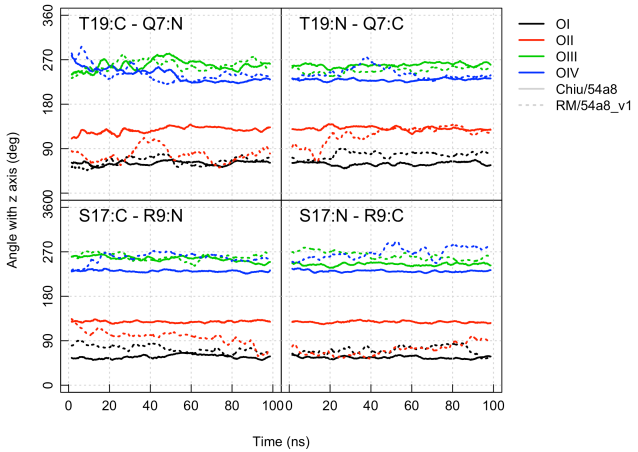


Figure 30: Angle with respect to the z-axis formed by the β-sheet plane. The backbone of amino acids 19 and 17 form hydrogen bonds with the backbone of amino acids 7 and 9 respectively. The vectors connecting the heavy atoms bonded to the O-H pairs forming them are taken as reference to compute the orientation. Solid lines refer to parameter set Chiu/54a8 and dashed ones to set RM/54a8_v1.

References

[1] M J Abraham et al. *GROMACS User Manual version 2016*. www.gromacs.org, 2018.

[2] S.-W. Chiu et al. "Incorporation of surface tension into molecular dynamics simulation of an interface: a fluid phase lipid bilayer membrane." In: *Biophysical journal* 69.4 (1995), pp. 1230–45. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(95)80005-6.

[3] R J Clarke. "The dipole potential of phospholipid membranes and methods for its detection". In: *Advances in Colloid and Interface Science* 89-90 (2001), pp. 263–281. ISSN: 0001-8686. DOI: 10.1016/S0001-8686(00)00061-0.

[4] W Ding et al. "Effects of Lipid Composition on Bilayer Membranes Quantified by All-Atom Molecular Dynamics". In: *The Journal of Physical Chemistry B* 119.49 (2015), pp. 15263–15274. ISSN: 1520-6106. DOI: 10.1021/acs.jpcb.5b06604.

[5] J P Douliez, A Léonard, and E J Dufourc. "Restatement of order parameters in biomembranes: calculation of C-C bond order parameters from C-D quadrupolar splittings." In: *Biophysical journal* 68.5 (1995), pp. 1727–39. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(95)80350-4.

[6] U Essmann et al. "A smooth particle mesh Ewald method". In: *The Journal of Chemical Physics* 103.19 (1995), pp. 8577–8593. ISSN: 0021-9606. DOI: 10.1063/1.470117.

[7] A A Gurtovenko and I Vattulainen. "Calculation of the electrostatic potential of lipid bilayers from molecular dynamics simulations: Methodological issues". In: *The Journal of Chemical Physics* 130.21 (2009), p. 215107. ISSN: 0021-9606. DOI: 10.1063/1.3148885.

[8] C Margreitter, M M Reif, and C Oostenbrink. "Update on phosphate and charged post-translationally modified amino acid parameters in the GROMOS force field". In: *Journal of Computational Chemistry* 38.10 (2017), pp. 714–720. ISSN: 01928651. DOI: 10.1002/jcc.24733.

[9] J F Nagle and S Tristram-Nagle. "Structure of lipid bilayers." In: *Biochimica et biophysica acta* 1469.3 (2000), pp. 159–95. ISSN: 0006-3002.

[10] C Oostenbrink et al. “A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6”. In: *Journal of Computational Chemistry* 25.13 (2004), pp. 1656–1676. ISSN: 1096-987X. DOI: 10.1002/jcc.20090. URL: http://dx.doi.org/10.1002/jcc.20090.

[11] J Pan et al. “Molecular structures of fluid phase phosphatidylglycerol bilayers as determined by small angle neutron and X-ray scattering”. In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1818.9 (2012), pp. 2135–2148. ISSN: 00052736. DOI: 10.1016/j.bbamem.2012.05.007.

[12] H I Petrache, S W Dodd, and M F Brown. “Area per lipid and acyl length distributions in fluid phosphatidylcholines determined by (2)H NMR spectroscopy.” In: *Biophysical journal* 79.6 (2000), pp. 3172–92. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(00)76551-9.

[13] A D Pickar and R Benz. “Transport of oppositely charged lipophilic probe ions in lipid bilayer membranes having various structures”. In: *The Journal of Membrane Biology* 44.3-4 (1978), pp. 353–376. ISSN: 0022-2631. DOI: 10.1007/BF01944229.

[14] T J Piggot, D A Holdbrook, and S Khalid. “Electroporation of the E. coli and S. Aureus Membranes: Molecular Dynamics Simulations of Complex Bacterial Membranes”. In: *The Journal of Physical Chemistry B* 115.45 (2011), pp. 13381–13388. ISSN: 1520-6106. DOI: 10.1021/jp207013v.

[15] D Poger and A E Mark. “On the Validation of Molecular Dynamics Simulations of Saturated and cis -Monounsaturated Phosphatidylcholine Lipid Bilayers: A Comparison with Experiment”. In: *Journal of Chemical Theory and Computation* 6.1 (2010), pp. 325–336. ISSN: 1549-9618. DOI: 10.1021/ct900487a.

[16] D Poger, W F Van Gunsteren, and A E Mark. “A new force field for simulating phosphatidylcholine bilayers”. In: *Journal of Computational Chemistry* 31.6 (2010), pp. 1117–1125. ISSN: 01928651. DOI: 10.1002/jcc.21396.

[17] R P Rand et al. “Variation in hydration forces between neutral phospholipid bilayers: evidence for hydration attraction”. In: *Biochemistry* 27.20 (1988), pp. 7711–7722. ISSN: 0006-2960. DOI: 10.1021/bi00420a021.

[18] M M Reif, P H Hünenberger, and C Oostenbrink. “New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field”. In: *Journal of Chemical Theory and Computation* 8.10 (2012), pp. 3705–3723. ISSN: 1549-9618. DOI: 10.1021/ct300156h.

[19] J Schamberger and R J Clarke. “Hydrophobic Ion Hydration and the Magnitude of the Dipole Potential”. In: *Biophysical Journal* 82.6 (2002), pp. 3081–3088. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(02)75649-X.

[20] T Starke-Peterkovic and R J Clarke. “Effect of headgroup on the dipole potential of phospholipid vesicles”. In: *European Biophysics Journal* 39.1 (2009), pp. 103–110. ISSN: 0175-7571. DOI: 10.1007/s00249-008-0392-y.

[21] I G Tironi et al. “A generalized reaction field method for molecular dynamics simulations”. In: 102.13 (1995), pp. 5451–5459. ISSN: 0021-9606. DOI: 10.1063/1.469273.

[22] L Wang, P S Bose, and F J Sigworth. “Using cryo-EM to measure the dipole potential of a lipid membrane”. In: *Proceedings of the National Academy of Sciences* 103.49 (2006), pp. 18528–18533. ISSN: 0027-8424. DOI: 10.1073/PNAS.0608714103.

[23] D T Warshaviak, M J Muellner, and M Chachisvilis. “Effect of membrane tension on the electric field and dipole potential of lipid bilayer membrane”. In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1808.10 (2011), pp. 2608–2617. ISSN: 00052736. DOI: 10.1016/j.bbamem.2011.06.010.

4.1 Additional material: effects of membrane undulation

During the revision process of the paper which constitute this Chapter, it was brought to our attention that membrane undulations are incompatible with the computation of ApL performed from the sides of the simulation box. The more a membrane patch is deviating from a flat geometry, the more approximate would be such method. Moreover, as the electron density profile is computed as an average over all the x and y positions, undulations of the membrane would result in a broader profile, with more uncertainty on the exact positions of the electron density peaks.

A brief comment on this appears on Section 2.5.1 of the paper, granting that the procedure chosen was suitable for the analysis of the simulations presented, and we would like to provide additional evidence to strengthen the case. This analysis supports as well the choice of using the box sides approach in Chapter 3 for simulation without or with low electric field, switching to a computation which takes into account undulations only for the highly curved membranes obtained with a high electric field.

The computation of ApL accounting for the membrane curvature is performed using a Fourier approach according to the work of Braun et al. [2011] and using the software made available from the same publication. This allows to obtain the ratio between the "true" area per lipid (i.e. computed taking into account the undulations of the membrane) and the projected one (i.e. from the box sides approach). Figure 4.1 reports the ratio $\mathrm{ApL_{und}}/\mathrm{ApL_{proj}}$ for each lipid and each parameter set: the values span between 1.0020 and 1.0046, which translates into an ApL correction between 0.20 and 0.46%, or, in absolute value, 0.001-0.003 nm^2. These discrepancies are lower than the standard deviation for any of the area per lipid computed, showing that the more accurate computation has no consequence for the purpose of evaluating the ApL and comparing parameter sets. The parameter set showing less discrepancy is Chiu/54a7, compatible with the fact that on average it produces smaller area per lipid: a more packed bilayer results in less freedom for the lipids to undulate.

Regarding the electron density, to check the influence of the averaging procedure over different locations of the patch, we computed the electron density for the phosphorus atoms on the full patch (256 lipid per leaflet), on a medium patch of roughly

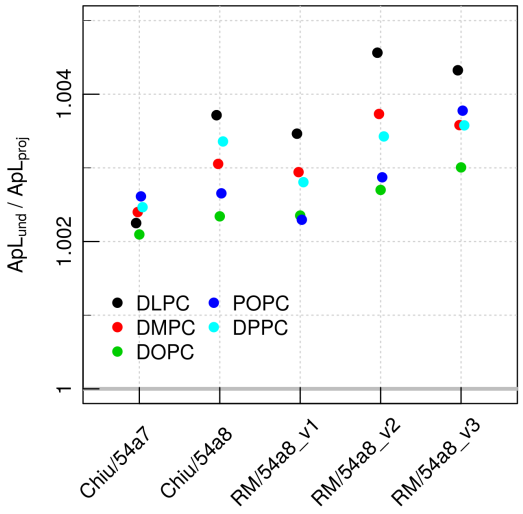


Figure 4.1: Ratio between area per lipid computed taking into account the undulations of the membrane (according to the procedure devised in Braun et al. [2011]) and the one computed from the projection on the plane parallel to the membrane. A value of 1 would be obtained for a perfectly flat membrane.

one forth in dimension (60 lipids per leaflet, see Figure 4.2(b) for a scheme), and finally a small 12 lipids per leaflet patch. It must be noticed that the GROMACS software, while computing the density perpendicular to the bilayer plane (presently along the z axis), centres the patch analysed around its center of mass, at each time step. This holds for GROMACS versions 5 onward, an version 2016.3 was used for all the analyses presented.

The resulting densities from the three different patches described above are normalised by the number of lipids present in each and compared. Figure 4.2(a) shows that the profiles for the full and medium patches are almost identical, while the smaller patch produces narrower peaks. Indeed, computing the full width at half maximum (FWHM) of these peaks, the FWHM of the small patch peaks is 16% smaller than the FWHM of the large patch peaks. However the distance between phosphorus peaks differ between patches by around 0.1 Å only, showing that the undulations have an effect mainly (or only) on the peaks broadness. As such, the hydrophobic thickness (measured as the distance between the phosphorous peaks) is virtually unvaried, and the difference imputable to the computing procedure is smaller than both the experimental thickness error and the differences arising from the use of different parameter sets. Further analysis must be done on the Luzzati thickness estimate, which depends

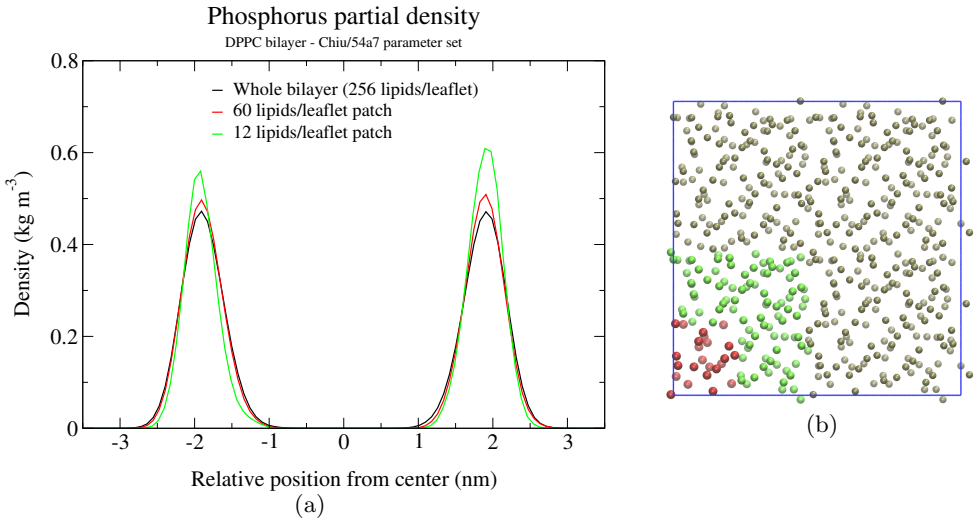


Figure 4.2: (a) Electron density of the phosphorus atom computed on the full membrane, a medium and a small patch. (b) van der Waals representation of the phosphorus atoms of the DPPC lipid bilayer (initial configuration) used for simulations. The simulation box is highlighted in blue. The green plus red regions constitute the medium patch mentioned in the analysis; the red one constitute the small patch.

on the broadness of the water profile. However, it must be noticed that also for this measure, the range of experimental results is quite broad and thus, as for many properties assessed in the paper, likely to be larger than the differences which can derive from the analysis procedure.

Another valid point raised in the discussion regards the use of the Berendsen barostat. As this algorithm does not reproduce the correct *isobaric ensemble* (see 2.1), the estimate of the lateral compressibility module is not accurate and should not be used to validate the parameters. Its computation is still valid as a comparison with the results from previous parametrisation works, but its relevance cannot be extended further.

Chapter 5

Conclusions

In this work we sought to understand and characterise at the molecular level the structures of a nanocapsule formed by antimicrobial peptides AMPs (named capzip) in solution. This assembly could not be determined by conventional structural investigations. In particular, we wanted to identify which amino acids or sub-structures play a role in the assembly to a nanocapsule unit and in its antimicrobial function. Multiscale simulations identified a network of β-sheets organised in a double layer as the preferred geometry for the assembly. The presence of a second layer was proved necessary to form a stable three dimensional object. This requirement allows to optimise the propensity of the residues in the sequence for solvent exposure by burying hydrophobic residues and exposing charged residues necessary to the antimicrobial activity.

To host both the hydrophilic and hydrophobic moieties present in AMPs, rigid α-helical structures have been in the past widely employed as they can accommodate both of them in a suitable way. However, the design of the peptide capzip explored the possibility of forming AMP structures without such fold. Our simulations have explained how the assembly of capzip is compatible with an amphiphatic structure formed by β-sheet peptides, once many copies are assembled together in a specific manner. The resulting double layer capsule successfully buries Tryptophan residues, which are the drivers of the stability and cohesion between the layers, and expose Arginine ones, which contribute to the β-sheet pairing. To pinpoint the role of specific amino acids was one of the aim of the work and it is particularly important in the context of designing new AMPs, as indeed a possible future direction of this work is the exploration of different amino acids composition for the molecule. This can hardly be done in a systematic

manner both at the experimental level and *in silico*, even employing coarse-grained simulations. Thus, ultra coarse-grained models could be employed in which capzip is represented with three/four beads for each strand. This would provide an important screening to understand which electrostatic and hydrophobic properties each segment should have to tailor the assembly.

The simulations of pre-assembled structures gave insight into the driving forces to assembly stability, particularly in comparing different coarse-grained force fields. One question which remains open concerns details on the kinetics of the process.

The second question this work aimed to answer was how the antimicrobial activity of the nanocapsule was performed, understanding which moieties actively perturbed the membrane and how such perturbation could be quantified. We used model membranes matching different experimental compositions, namely a bacterial membrane composed by zwitterionic and anionic phospholipids, and a mammalian one composed only by zwitterionic phospholipids. We assessed that capzip interacts favourably with the bacterial membrane, but less favourably with the mammalian one. This property is needed to selectively exploit the antimicrobial activity. The peptide decreases the stability of the bacterial membrane and its ability to withstand external perturbation like the one of an electric field. This observations illustrate how capzip can promote poration of the bacterial membranes only.

The disruptive action is lead by charged Arginine insertion in the hydrophobic membrane region, while we do not observe the hydrophobic residues (Tryptophan) to penetrate deeply into it. This is in disagreement with the generally accepted picture of antimicrobial peptide insertion into membranes, which proposes a two step process formed by a charge binding event followed by hydrophobic insertion. However, this mode of action was postulated for helical AMPs. Here we proved instead that for an assembly of AMPs with high β-sheet content, the Arginine insertion is sufficient to trigger poration without rearrangement of all the Tryptophan within the membrane by applying an electric field close to physiological values. The assembly of capzip is crucial in bringing many antimicrobial sequences within the same region of the membrane, creating *loci* with high positive charge. Moreover, through the presence of many Arginine residues, a chain of water molecules inserting into the hydrophobic membrane

core triggers the formation of a pore, that leads to the disruption of the membrane.

This investigation necessarily explores specific, simplified membranes. The results correlate with the experiments on Supported Lipid Bilayer of the same composition, informing the mechanism that allow their disruption. However, they must be reconciled with the complex structure of both bacterial and mammalian real membranes: the bacterial model used is mimicking the inner membrane, and therefore the mechanism of permeation and/or transportation of the peptide through the outer membrane (for Gram-negative bacteria) or the peptidoglycan layer (for Gram-positive ones) remains still to be elucidated. However, the bacterial cell wall is negatively charged, as in the adopted model. Therefore we can assume that the mechanism of adsorption and disruption observed in our simulations is in broad sense reproducing the physiological situation.

On the contrary, the scarce interaction observed with the mammalian membrane poses the question of which determinants are then favouring a good internalisation of siRNA molecules in HeLa cells, as proven by experiments Castelletto et al. [2016]. The resulting system (siRNA plus capzip assembly) is very different with respect to the pure capzip assembly because of the negative charge contributed by siRNA molecules. As such, it has not been investigated in this work. The modelling of this is within the future outlook of this work, important for applications of capzip for delivery and alternative therapies. Indeed, the capsule itself has the potential to transport a broad variety of drugs, rather than siRNA only, but the characteristics of the capzip-drug assembly must be specifically modelled for each case to understand the interaction of the nanocapsule carrier with bacterial and mammalian membranes. These considerations make an exciting ground for boosting the application of capzip assembly not only as an antimicrobial agent but also as a delivery agent, fully exploiting the multi-functional behaviour it was designed for.

Finally, this work brought two contributions to methodological aspects of simulations techniques: they emerged as a valuable by-product of the *in silico* exploration chosen, in the effort of developing and applying the best and more updated techniques to our system.

First, connected with the exploration of capzip assembly, we performed a systematic

analysis of the performances of different coarse-grained force fields (SIRAH, MARTINI, Polar MARTINI) in simulating the assembly of the capzip. In particular we proved that the SIRAH force field returns an energetic profile closer to the atomistic one, while the MARTINI models proposed a remarkably different balance between the electrostatic and short range energetic components. These conclusions could be a result of the model underlying the respective parametrisations, as SIRAH groups fewer atoms than MARTINI in one bead, however, to our knowledge, this is the first comparison which presents a quantitative analysis of it. This is important for the choice of the coarse-grained model, in particular when simulating systems with a high net charge, as the relative weight of the Coulomb interactions versus short range ones may vary considerably.

Finally, we carefully analysed and questioned the parametrisation of lipids in the GROMOS force field, proposing a more up to date version of parameters for phopsholipids and phosphocholine in general (called version 54A8_v1, available at `http://fraternalilab.kcl.ac.uk/wordpress/biomembrane-simulations/`). This novel parametrisation has a significant impact on the simulations of proteins-lipids interaction, as proved by test simulations of the AMP Lfcin. We think that this parameters are a first step toward reconciling the description of the two components, to model their mutual influence at best. This work places itself in the workflow of continuous improvements of MD simulations parameters, which is constantly prompted by three factors: better experimental results available, the development of new strategies of parametrisation, and the evolution of computers, which allows the simulations of larger systems and longer time scales, verifying or questioning what already achieved exploring reduced systems or shorter times. The past successes of MD simulations already show the potential of this techniques which nevertheless must be continuously improved to tackle more complex systems. As such, we are keen in testing more extensively the new parameters and to repeat selected simulations of capzip on model membranes to understand the type of interactions that they propose.

Overall, this work elucidate some principles and rules in the assembly and antimicrobial mechanism of actions of a synthetic molecule by use of computational techniques: these processes are inaccessible to experiments, proving that Molecular Simulations are

an irreplaceable tool to access their atomistic details.

Appendices

A.1 Additional publications

The following pages present the manuscript: Kepiro, I. E., Marzuoli, I., Hammond, K., et al. Engineering chirally blind protein pseudo-capsids into nanoprecise antibacterial persisters, submitted and in second revision for ACS Nano. Chapter 3 includes part of the results presented in the paper, elaborating them further.

Engineering persistent antimicrobial pseudo-capsids

Ibolya E Kepiro,[1] Irene Marzuoli,[1,2] Katharine Hammond,[1, 3, 4] Xiaoliang Ba,[5] Helen Lewis,[1] Michael Shaw,[1] Smita B Gunnoo,[1] Emiliana De Santis,[1] Urszula Lapinska,[6] Stefano Pagliara,[6] Mark Holmes,[5] Christian D Lorenz,[7] Bart W Hoogenboom,[3] Franca Fraternali[2] and Maxim G Ryadnov[1,]*

[1]National Physical Laboratory, Hampton Road, Teddington, TW11 0LW, UK

[2]Randall Centre for Cell and Molecular Biophysics, King's College London, London, SE1 1UL, UK

[3]Department of Physics and Astronomy, University College London, London, WC1E 6BT, UK

[4]London Centre for Nanotechnology, University College London, London, WC1H 0AH, UK

[5]Department of Veterinary Medicine, University of Cambridge, Cambridge, CB3 0ES, UK

[6]Living Systems Institute, University of Exeter, Exeter, EX4 4QD, UK

[7]Department of Physics, King's College London, Strand Lane, London WC2R 2LS, UK

Corresponding author:

Dr Maxim G Ryadnov

National Physical Laboratory,

Hampton Road, Teddington, TW11 0LW, UK

Fax: (+44) 20 86140573

Tel: (+44) 20 89436078

max.ryadnov@npl.co.uk

ABSTRACT: the spread of bacterial strains persistent to antibiotic treatments stimulates the search for an alternative means of antimicrobial interventions. Here we report a virus-like pseudo-capsid engineered to self-assemble from re-purposed effector domains of a host defence protein to exhibit a broad spectrum of antibacterial activities in vitro and in vivo. Using a combination of nanoscale and single-cell imaging we demonstrate that such pseudo-capsids inflict instantaneous and all-out damage to bacterial cells by penetrating and disrupting their membranes. Unlike conventional antibiotics, these agents are effective against phenotypic bacterial variants, while clearing superbugs in vivo without toxicity. The monomeric subunits of the designed pseudo-capsids can be in either epimeric form, L or D, providing a versatile platform for engineering structurally diverse and functionally persistent antimicrobial capsids.

Antimicrobial resistance is a natural process that helps bacteria adapt to new environments (*1*). The process is relentless and represents a major health threat given our dependence on antimicrobial treatments (*2*). With conventional antibiotics losing their effectivity at an alarming rate, alternative antimicrobial interventions attract an ever-increasing interest (*3*). Most compounds that are prescribed to patients today act by binding to individual intracellular targets in bacterial cells. A single genetic event is sufficient for bacteria to acquire potent counter measures. The time these take to manifest can fall within just a few years after a given antibiotic is introduced to the market (*3*). As a consequence, bacteria can develop into "superbugs" – strains that no longer respond to antibiotic treatments. The spread of these pathogens necessitates the development of novel antimicrobials that can rapidly attack a bacterial cell as a whole. This strategy may provide agents capable of killing growing, persister and dormant cells – this feat is inaccessible to conventional antibiotics (*4*). New agents should be able to differentiate between host and bacterial cells as the attack is likely to involve cell membranes. Strictly speaking, the innate immune systems of multicellular organisms do not use antibiotics. Instead, they deploy host-defence effector molecules to the sites of infection. These molecules are stand-alone

peptides or relatively small domains in globular proteins that recognise pathogen associated molecular patterns on the surfaces of microbial cells (*5*). These effector molecules are diverse in structure and origin, but share common physicochemical properties. Most of them are cationic and fold into amphipathic conformations upon binding to anionic microbial membranes (*6*). In membranes, these conformations assemble into higher-order oligomers that overcome a threshold of peptide concentration, beyond which antimicrobial effects become apparent. Reaching this concentration appears to be a limiting step in the ability of the oligomers to permeate membranes (*7*), and is more characteristic of host defence peptides, e.g. human defensins and cathelicidins (*8*), than of effector domains of multifunctional proteins, e.g. lactoferrins, that recognise pathogen surfaces without the need to self-oligomerize (*9*). Nonetheless, in free, isolated forms, these domains can oligomerize to induce bacteriostatic effects (*10*). This presents a considerable opportunity for antimicrobial engineering. Multiple copies of an isolated effector domain can be arranged to fold into a discrete and autonomous assembly. The domain copies of the assembly are ready to be deployed at the site of contact with microbial membranes, where they instantaneously deliver concentrations that significantly exceed those necessary to rupture microbial membranes. An ultimate benefit of this design is rapid and all-out damage to a microbial cell without the need for the transition from unstructured peptide monomers to membrane-active oligomers – a strategy that may be less subject to acquired resistance. However, the success of this approach relies on the ability of the effector molecules to pre-concentrate in a distinct nanoscale structure that can bind to microbial membranes.

More specifically, most bacterial cells do not exceed 0.2 to 1 μm in width. Therefore, a discrete assembly of effector domains ranging from tens to hundreds of nanometres in diameter would be sufficient to inflict irreparable damage to a bacterial cell. To be discrete at these size ranges the assembly is best confined to a platonic, symmetrical structure. The nature of the symmetry is of less importance as long as the assembly is locked into a three-dimensional form that can bind to microbial membranes. In this regard, viral capsids inspire a straightforward solution. These are self-assembled

nanoscale shells that can be commensalistic to multicellular hosts. Viruses do not kill bacteria on contact, but can provide suitable architectural templates for the capsid-like assembly of effector molecules. Resulting assemblies need not be strictly monodisperse because their designed function allows for a level of structural polymorphism that is typical of aberrant and pseudo-capsids (*11-13*). Furthermore, the assigned function imposes no *a priori* constraints on the chirality of the effector domains. Polypeptide chains of reversed chirality or all-D peptides are much more stable to proteolysis, fold in a similar manner to that of all-L peptides and are likely to be non-immunogenic (*14-17*). Antimicrobial D epimers are as effective as their L counterparts, which is consistent with that host defence peptides act by binding to the lipid components of bacterial membranes rather than dock to a specific protein (*17, 18*). Herein we demonstrate the application of the outlined principles for the design of antimicrobial pseudo-capsids or Ψ-capsids.

RESULTS

Pseudo-capsid design

Our approach adapts a short host defence motif from a multifunctional protein lactoferrin – a major component of the innate immune system (*9, 10*). The antimicrobial properties of this protein are attributed to its N-terminal domain. The domain has a broad spectrum of targets including porins, DNA, intracellular metabolites, and can stimulate the immune system by neutralising endotoxin (*19-21*). The host defence motif of the domain is a hexapeptide RRWQWR, which has strong propensity for antiparallel β-sheet conformations (*22*).

To render this motif self-complementary, the C-terminal arginine of the peptide was replaced with a glutamate. This modification facilitates inter-strand Coulombic interactions with the N-terminal arginine of the opposite β-strand. The core of the motif is also homologous to a characteristic motif of tryptophan zippers, WTW, which folds with cross-strand tryptophan rings packed tightly against one another (*23*). To capitalise on this analogy, the glutamine in the peptide was replaced with a threonyl residue. This modification supports the formation of a tryptophanyl interface to cement a β-sheet

bilayer with two cationic exteriors. In viruses, capsid proteins interface orthogonally with each other to network into penta- and hexagonal assembly units. Each of these units adopts a three-fold rotational symmetry which ensures their propagation into a closed symmetry, shell (*24*). To emulate the three-fold symmetry of native shell-like subunits, the resulting motif RRWTWE was converted into a triskel conjugate (Figs 1A and S1 in Supporting Information), with both L- and D-forms of this conjugate chemically synthesised (Fig S2).

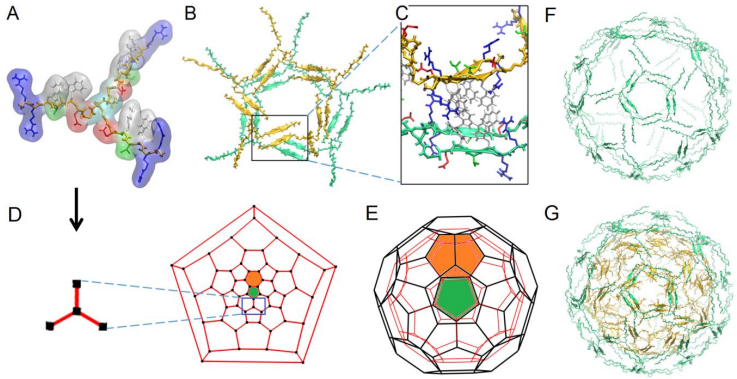


Figure 1. Pseudo-capsid design. (A) A molecular model of the triskel RRWTWE conjugate. (B) A snapshot of molecular dynamics simulations showing a pentagonal assembly unit formed by designed triskelions. (C) Two β-sheet arms (green and yellow) forming a bilayer interface via cross-strand packed tryptophans (grey). (D) A schematic representation of the designed triskelion as a monomer subunit in a truncated icosahedron shown as a 5-fold Schlegel orthographic projection. (E) 5-fold symmetry orthographic projections of two truncated icosahedra (black and red) forming a double-walled assembly. Note: for clarity only one of each of the pentagonal (green) and hexagonal (orange) units is highlighted in D and E. (F) A single-walled triskelion assembly templated on a truncated icosahedron. (G) A double-walled triskelion assembly templated on a truncated icosahedron. Outer and inner shells in (F) and (G) are shown in green and yellow, respectively.

Each arm of this triskelion pairs into an antiparallel β-sheet with another arm of another triskelion. Each folded pair of two arms interfaces with another folded pair of two other conjugates (Fig 1B, C). The monomer has a tri-lateral symmetry which enables it to assemble into penta- and hexagonal units forming β-sheet bilayer networks (Fig 1B). Because β-strands are stabilised by *inter*-strand interactions it is physically impossible for the bilayers to have free, "sticky" edges. These have to close on one another. Triskelion structures have an intrinsic non-zero curvature (*25-27*). In propagating β-sheet networks, the curvature translates into a trans-sheet asymmetry prompting the spontaneous closure of increasingly curved sheets into a minimum energy structure – shell (*27, 28*). This mode of assembly is analogous to that of viral shells which follows an icosahedral symmetry, with a truncated icosahedron being one of the most common architectures (*29, 30*). This polygon architecture offers an ideal template for the assembly of triskelions: it has only three-fold vertices, each of which can host an individual triskelion (Fig 1D, E). Therefore, the cooperative assembly of triskelions within the template should create an equilibrated β-sheet shell. The monolayer configuration of this structure is not stable as the hydrophobic side chains of the tryptophanyl residues are oriented inwards towards the water-filled core of the shell (Fig 1F). A β-sheet bilayer or a double-wall shell, in which cationic arginyl residues furnish its exterior and interior surfaces, is deemed more stable (Fig 1C, G) (*31*). In support of this conjecture, coarse grain molecular dynamics (MD) simulations (*32*) showed that single-wall icosahedra assembled from the triskelion collapsed within the first 600 ns of simulations following an equilibration phase (Fig S3A). Under the same simulation conditions, double-walled shells retained their initial configuration over 1 μs of coarse grain simulations, evolving into more compact shapes, for both L- and D-forms, which was also confirmed by 100 ns atomistic simulations (*33*) (Figs S3B, S4 and Movie S1). The results of the simulations indicate that triskelions can assemble cooperatively in the template reaching an equilibrated and stable structure.

Pseudo-capsid folding and assembly

Consistent with the simulations, Circular Dichroism (CD) spectra for both forms of the triskelion were characteristic of antiparallel β-sheet and β-turn conformations (Fig S5A). Fourier transform infrared (FT-IR) spectra revealed bands for the de-convoluted amide regions at 1650-1670 cm^{-1} for β-turn structures and at 1630 cm^{-1} and 1545 cm^{-1} for β-sheets (Fig S5B).

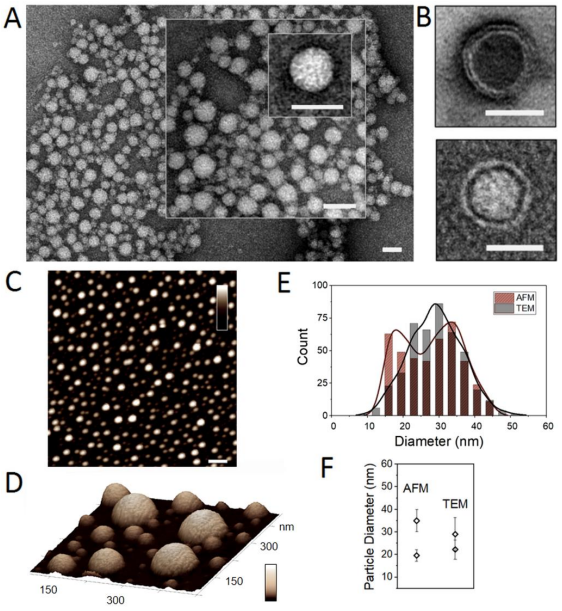


Figure 2. Pseudo-capsid assembly. (A) TEM images of assembled D-Ψ-capsids. Scale bars are 50 nm. (B) Higher resolution TEM images of individual collapsed capsids. Scale bars are 50 nm. (C) Topography images of D-Ψ-capsids obtained on a mica substrate by in liquid AFM. Colour (height) and scale bars are 60 nm and 200 nm, respectively. (D) 3D representation of D-Ψ-capsids. Colour (height) bar is 65 nm. (E) Size distributions and dominating sizes of D-Ψ-capsids by AFM and TEM. (F) Average sizes of dominating populations of Ψ-capsids. Assembly conditions: 100 μM peptide, pH 7.4, 10 mM MOPS, 20˚C, overnight.

The results suggest that inter- and intra-molecular hydrogen bonds in β-sheets and β-turns, respectively, support the cooperative folding of the triskelions into higher-order structures. Indeed, transmission electron microscopy (TEM) and atomic force microscopy (AFM), performed in solution, revealed uniformly spherical capsid-like shells (Fig 2A-D), some of which appeared to collapse into a double-walled or double-layered morphology (Fig 2B). The diameters of the shells were consistent by width (TEM) and height (AFM) measurements showing narrow step-size distributions, e.g. 10-20 nm, 20-40 nm (Fig 2E, F). This suggests that the triskelions may adopt an integer step size in assembly, generating shells that are polymorphic in size, but not in shape. This property is akin to that of viral capsids and synthetic virus-like assemblies that can adapt to repack into smaller and larger shells (*34-36*). This is also in marked contrast to mutant viral capsids and polymorphic virus-like particles whose variations in morphology feature filamentous and aberrant structures with irregular serrations (*12*, *13*). Taken together the findings indicate that the designed triskelions propagate with the formation of thermodynamically stable Ψ-capsids exhibiting a degree of structural plasticity that helps accommodate size variations without compromising on morphological uniformity. Since the triskelions occupy the vertices of pentagonal and hexagonal faces in the truncated icosahedron, rather than tightly pack in the faces, the overall architecture of Ψ-capsids remains independent of size allowing for wide size variations. This property renders the assembly adaptable to morphological changes imposed by hierarchically complex and dynamic environments such as microbial membranes. To test their behaviour in membranes in sufficient detail, Ψ-capsids were introduced into reconstituted phospholipid bilayers that were assembled on mica substrates as described elsewhere (*37*). The resulting supported lipid bilayers (SLBs) provide suitable models for bacterial membranes, are flat (to within ~0.1 nm) in their unperturbed state (Fig 3A, B), and allow for the accurate depth measurements of surface changes in solution and in real time by AFM (*38*). As gauged by these measurements, Ψ-capsids bound to the SLBs and disintegrated on them forming pore- and channel-like lesions (Fig 3A-C).

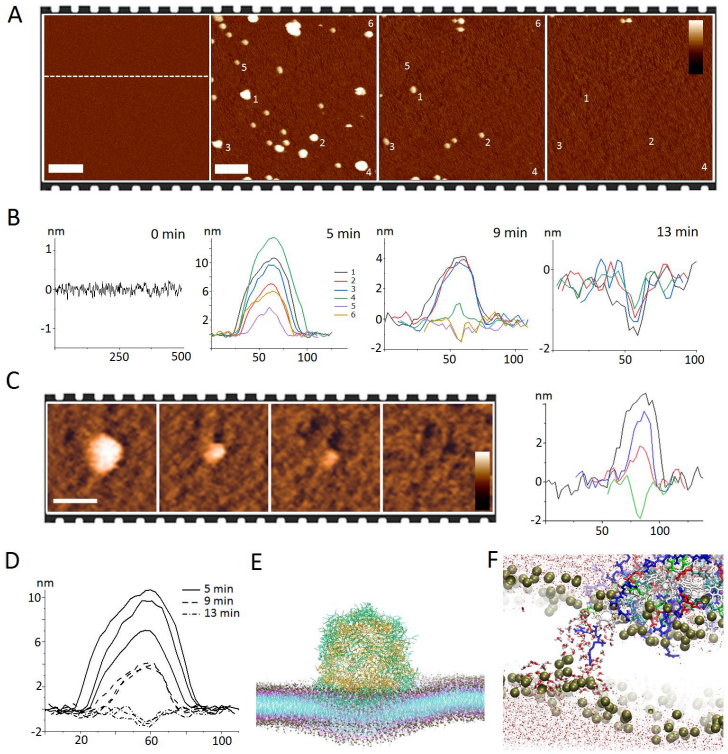


Figure 3. Pseudo-capsids porating phospholipid bilayers. (A) In liquid AFM topography of SLBs treated with D/Ψ-capsids. The images are taken at 4-min intervals. Individual Ψ-capsids are numbered (1-6) to highlight conversions into pores. Scale bars are 100 nm for the first image and 200 nm for the others. The height (colour) bar is 10 nm. (B) Cross-sectional analyses of Ψ-capsids numbered as in (A) for each given time point. (C) Topography of a Ψ-capsid converting into a pore in 15 min, with the height profiles of the corresponding conversion area. Images are taken at 5-min intervals. Scale and height (colour) bars are 50 nm and 10 nm, respectively. (D) A plot combining height profiles of three Ψ-capsids from (A) to show conversions as a function

of time. (E) A snapshot of coarse-grain MD simulations of a Ψ-capsid landing on a phospholipid bilayer at 2 μs and simulated for additional 8 μs. Key: outer and inner shells are in green and yellow, respectively; golden beads denote phosphate groups linked to two types of polar groups (blue and magenta); aliphatic lipid tails are in cyan. (F) A 72-ns snapshot from 100-ns atomistic simulations of a rudimentary pore showing a triskelion arm stretching across the bilayer interface. Key: for clarity, only phosphate groups (golden beads) are shown for the bilayer.

The conversion of individual shells into pores at their precise landing positions was complete within a few minutes of treatment: Ψ-capsids gradually sank in the lipid bilayers at the depths of a folded triskelion arm, ~1.8 nm (0.3 nm translation per β-strand residues) (Fig 3A-D). Coarse grain simulations of the capsids in the lipid bilayers (*39*) run over 10 μs revealed that the sinking effect occurred at the expense of pressing and displacing lipids deep into the bilayer interface (Fig 3E, Movie S2). In accord with this, atomistic simulations of a rudimentary pore showed that triskelion arms oriented towards the bilayer interface (Figs 3F, S6 and Movie S3). Such localised responses are consistent with a mechanism where triskelions re-assemble in the bilayer and that forming pores do not expand and remain confined within the diameters of Ψ-capsids. This is important for three reasons. First of all, the assemblies exhibit a larger structural plasticity than crystalline materials or viruses do, which allows them to rapidly re-arrange into peptide-lipid oligomers at the sites of contact with phospholipid membranes. Secondly, these oligomers maintain the size of the resulting pores close to that of the landed capsids thereby supporting precise and site-specific membrane disruption. Thirdly, this behaviour suggests that Ψ-capsids may support differential responses in cell environments favouring attack on microbial membranes. To gain a better insight into this, the biological properties of Ψ-capsids were assessed using a range of *in vitro* and *in vivo* assays.

Biological properties of pseudo-capsids

Irrespective of chirality, Ψ-capsids were found to be antimicrobial and non-hemolytic (Table S1). The minimum inhibitory concentrations (MICs) of the capsids were comparable to those of antimicrobial

agents including conventional antibiotics (Table S1). However, the principal advantage of Ψ-capsids over antimicrobial compounds lies in their ability to exert rapid and all-out damage to a bacterial cell, which makes them equally effective against susceptible and tolerant cells. In liquid AFM experiments showed that the capsids can indeed disrupt membranes within minutes. MIC experiments cannot directly relate to the AFM results as these are optical density measurements that do not take into account changes at the cellular level. Therefore, we sought complementary evidence from three series of experiments using planktonic and sessile bacterial culture of two of the most common pathogens – *E. coli* and *S. aureus*. In the first series, the antimicrobial activity of Ψ-capsids was assessed as a measure of total cell counts following capsid treatments (*27*). Negligible cell counts were observed for cultures treated with the capsids in comparison to the samples of untreated cells in which appreciable bacterial growth was observed (Fig S7). These findings provide end-point results of treated cell populations obtained over the same timescale as MIC measurements. To elucidate the antimicrobial kinetics of Ψ-capsids within the first hours and with a single-cell resolution, the second series of experiments was conducted. In these experiments thousands of individual *E. coli* cells were screened using a high throughput microfluidic device comprising thousands of growth channels (*40*). Each channel serves as a trap for an individual cell where the cell can grow through one of its distal poles, whereas the diameter of the channel matches that of the cell thereby arresting its movement (*41*). The replicative age of this trapped, mother, cell increases by one generation at each cell division allowing to follow phenotypic cell inheritance for an infinite number of generations. Such a "mother machine" then enables the single-cell monitoring of antimicrobial kinetics *in situ* and is able to reflect the responses of different cell phenotypes to antimicrobial agents (*42*). With this in mind, Ψ-capsids were introduced into *E. coli* cells trapped in the channels of the device (Fig 4A). After three hours of treatment under a constant flow the microfluidic environment was replaced by flowing the culture medium over 21 hours to restore normal conditions for bacterial growth. Subsequent analyses revealed

that Ψ-capsids killed all encountered cells (Fig 4). By contrast, persister cells and viable but not-culturable (VBNC) cells were found after ampicillin treatments (Fig 4A, B).

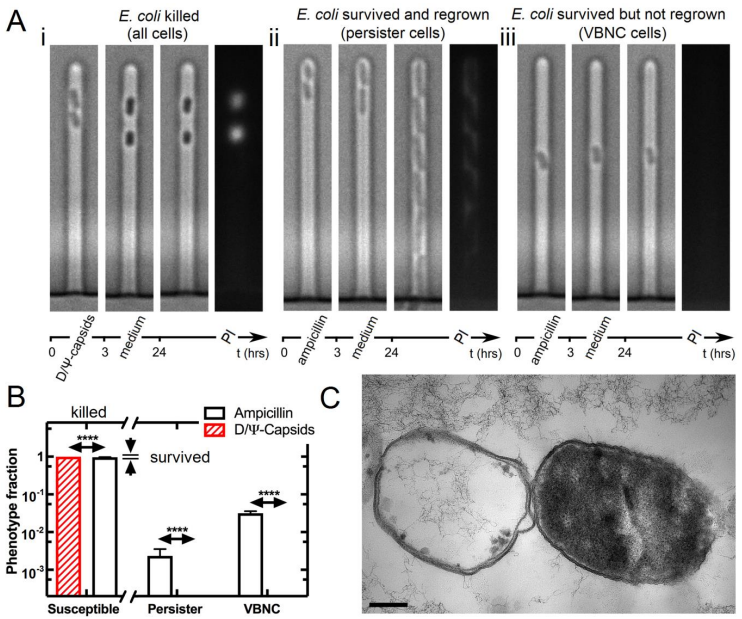


Figure 4. Single-cell antimicrobial kinetics of pseudo-capsids. (A) Three panels of representative optical micrographs of *E. coli* cells (i) killed by D/Ψ-capsids; (ii) survived and regrown after ampicillin treatments (persister cells); (iii) survived but not re-grown after ampicillin treatments (VBNC cells). In each panel, image bars show (from left to right) cells after 3-hour treatments with capsids or ampicillin, during subsequent washing with the growth medium over 21 hours, and after washing. The last micrographs in the panels are fluorescence images of 24-hour samples stained with propidium iodide (PI) – a live-dead stain penetrating dead bacteria with compromised membranes. (B) Cell fractions treated with D/Ψ-capsids and ampicillin. The data and error bars are means and standard error of the means of measurements obtained for 3332 cells hosted in 2331 independent

microfluidic channels in four independent biological replicates. The distribution of fractions of susceptible cells to ampicillin was significantly lower than that to D/Ψ-capsids (p <0.0001). The distributions of fractions of persister and VBNC cells to ampicillin were significantly higher than those to D/Ψ-capsids (p <0.0001). (C) Electron micrographs of microtomed *E. coli* cells treated with D/Ψ-capsids. The scale bar is 200 nm.

These tolerant phenotypes represent common sub-populations in clonal bacterial cultures that persist antibiotic treatments even at high doses (*43*). Persister cells resume growth after the drug is removed from their environment, whereas the regrowth of VBNC cells often requires specific conditions (*44*, *45*). Both phenotypes contribute to infection relapses prompting repetitive treatments and can be linked to dormancy (*46*). With Ψ-capsids effectively reducing these phenotypes to susceptible cells, changes in cell morphology may shed light on the mechanism of action at the single cell level. Indeed, the capsid-treated cells appeared shrunk and somewhat denser when compared to unaffected cells suggesting that the capsids accumulated in the membranes. This is consistent with the AFM findings (Fig 3), though the cells in the mother machine did not disintegrate and seemed intact (Fig 4B). Therefore, cells treated with capsids were microtomed and imaged by TEM. Complete membrane destruction was evident (Figs S8). Intriguingly, however, many cells appeared as empty and half-empty carcases with displaced and bulging membranes that tended to open up at distal poles (Figs 4C, S8). This effect is not normally observed for bacteria damaged by host defence peptides or membrane-active antibiotics, which porate bacterial membranes indiscriminately, causing cells to shrink and leak (*37*) or, as polymyxins, aggregate with lipopolysaccharides into membrane-destabilising blebs across the whole bacterial surface (*47*). The emphasis on distal poles is reminiscent of bacteriophages that preferentially target the poles of bacterial cells (*48*). Although, like polymyxins, Ψ-capsids favor anionic lipids, morphologically they are more of phage mimetics that may be attracted to outward membrane curvature or cell poles serving as DNA mobilisation sites before cell division (*48*, *49*). The third series of experiments provided further insight into this. The attack of Ψ-capsids on *E. coli* cells was monitored by structured illumination microscopy (SIM), which was custom designed to image

biomolecular and cellular processes in situ (*50-52*). The capsids rapidly adhered to the cell surfaces, with adsorption at the distal poles being also apparent (Figs 5A and S9A). Within a doubling time (20 min), the capsids proved to come into direct contact with the cells accumulating in their membranes and cytoplasm (Fig 5B, C). The affected cells then underwent sharp rupture accompanied by a burst of fluorescence intensity (Figs 5C, S9A-C and Movie S4). This effect was more profound in cocci cells (*S. aureus*). Although the distal poles of these cells are less defined, the transition from the initial contact with Ψ-capsids to the disintegration of their only membrane was somewhat sharper when compared to that for *E. coli* (Figs 5C, S9B and Movie S5).

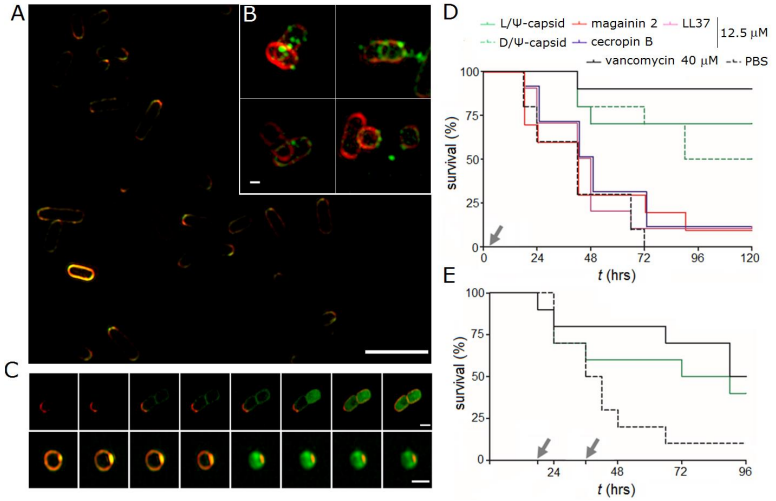


Figure 5. Time-kill kinetics of pseudo-capsids. Structured illumination micrographs of *E. coli* cells (A) immediately after the addition of L/Ψ-capsids and (B) during the treatment highlighting individual capsids in contact with cells. The capsids are labelled with carboxyfluorescein (green). Cell membranes are stained with FM4-64FX (red). Scale bars are 5 μm and 500 nm for (A) and (B), respectively. (C) SIM images recorded at 5-min intervals for *E. coli* (upper) and *S. aureus* (lower) incubated with the capsids. Scale bars are 1 μm. (D) Survival of *G. mellonella* larvae infected with vancomycin-susceptible MRSA NCTC12493 strain when treated

with Ψ-capsids and host defence peptides. Vancomycin and phosphate buffered saline (PBS) were used as positive and negative controls, respectively. Inoculations were done straight after the initiation of infection (first 2 hrs) without subsequent treatments. (E) Survival of MRSA-infected *G. mellonella* larvae treated with vancomycin, L/Ψ-capsids and PBS, administered at 18 and 36 hrs after the initiation of infection. Survival rates for Ψ-capsids were significantly higher when compared to PBS control (Mantel-Cox test, $p < 0.001$). Grey arrows indicate inoculation time points.

Strikingly, the time of cell disruption closely matched that needed for the conversion of the capsids into membrane-disrupting pores as gauged by AFM (Fig 3). Collectively, these findings suggest that Ψ-capsids need not disintegrate in the membranes, which they may effectively traverse reaching the protoplasm. Such a mode of action infers that the capsids may be able to circulate over prolonged periods of time in cellular environments that are challenged by sustained and active bacterial growth. To verify this in a biologically relevant model, Ψ-capsids were administered into *G. mellonella* larvae infected with a methicillin-resistant *S. aureus* (MRSA) strain susceptible to vancomycin (*53*, *54*). Over the first 48 hours the treatments were as effective as those by high doses of vancomycin (Fig 5D). The survival rate of up to 80% proved to be steady for another 24 hours for larvae treated with D/Ψ-capsids, whereas a 10% drop was observed for larvae treated with the L-form (Fig 5D). The resulting rate of 70% did not change over the total of 120 hours of treatment for L/Ψ-capsids. In comparison, for the D-form it dropped to 50% on day 4, suggesting a depletion of the capsids in the larvae. The exact reason for this is unclear, though the decrease is likely to link to that D/Ψ-capsids were more active during the first 72 hours, possibly at the expense of more capsids being engaged with bacteria (Fig 5D). The observed survival rates proved to be superior over those reported for membrane-active antibiotics and bacteriocins implying that the capsids may indeed be equally effective in killing bacterial cells from the outside and inside (*54*-*56*). These experiments also demonstrated the importance of pre-assembling effector monomers into capsids, which was re-emphasised by the failure of stand-alone host-defence peptides to exert sustained antimicrobial effects under the same conditions.

Specifically, the administration of human (LL37), insect (cecropin B) and animal (magainin 2) antimicrobial peptides into the infected larvae did not improve survival rates relative to those of mock injections with PBS (Fig 5D). Furthermore, the capsids maintained high rates of bacterial clearance even when administered at significantly delayed injection times: the larvae treated with capsids at 18 and 36 hours exhibited a 50% survival up to day 4 post-infection (Fig 5E), which is twice the time achieved by bacteriophage therapies (*56*). Finally, multiple injections of the capsids into uninfected larvae did not lead to appreciable cytotoxic effects (Fig S10).

DISCUSSION

The obtained results prompt several conclusions. First of all, we have demonstrated that a discrete virus-like shell assembled from moderately antimicrobial effector molecules delivers instantaneous and irreparable damage to a bacterial cell. The damage results from a direct impact that the assembly has on the cell upon contact and ensues via multiple avenues starting with the rapid conversion of the assembled shells into membrane-disrupting pores. Remarkably, pore formation is spatially confined to the landing position of a shell, which ensures the rapid and highly localized influx of high antimicrobial doses at the site of contact. Such an impact is different from that of organic or inorganic nanoparticles that are able to rupture membranes indiscriminately by a purely physical means, are unable to disintegrate and are prone to agglomeration. In contrast, the shells are cooperative ensembles of monomeric units that are capable of re-assembly in bacterial membranes and subsequent integration into the protoplasm. Thus, secondly, these ensembles yield equilibrated nanoscale systems whose construction is symmetry driven. This is a prerequisite for the structures of this type if they are to remain autonomously operational within the nanoscale size range (*24*, *57*, *58*). In this regard, the polygonal architecture of viral capsids proved to be a suitable framework for templating the assembly of monomeric units, which can be of virtually any chirality and topology. An efficient strategy for the design is then to re-purpose host defence effector molecules into structurally and functionally tunable bactericidal motifs that are furnished with the ability to self-assemble. Individually, such motifs may

resemble membrane-active antibiotics or host-defence peptides, but their efficacy is no longer subject to the lag phase of reaching a critical threshold concentration on bacterial surfaces. Once assembled into a geometrically locked pseudo-capsid the motifs acquire a dual capacity of pre-concentrated and stimuli responsive antimicrobial doses. Their responsiveness builds upon their structural plasticity, which allows them to accommodate changes in dynamic environments without being restricted to a particular size (*27, 59*). Subsequently, requirements for nanoscale order are looser than those for viral subunits that tune their assembly to the size of encapsidated genomes. Similar to other protein paracrystals, which often exhibit periodic nanoscale patterns (e.g. striations or rings) (*60, 61*), the designed pseudo-capsids display fine surface structure (Fig 2A-D). Although this did not feature readily recognizable patterns, the assemblies appeared to adopt a size integer while being monomorphic in shape without competing morphologies such as filaments (*62*). Combined these physicochemical properties give rise to a highly effective antimicrobial system that, unlike antibiotics, is not frustrated by antibiotic tolerant phenotypes such as persister or VBNC cells, or by superbugs, killing all. This outcome is also notable in that the design may provide a useful tool to aid in a better understanding of how different bacteria phenotypes can be distinguished and selectively targeted (*63*). A question remains as to how universal the ability of such pseudo-capsids is in overcoming different resistance mechanisms including cell surface fortifications, efflux pump blockages or peptide antagonists that block access to cell membranes (*64-67*). As it has been long postulated (*68*) and re-emphasized most recently (*69*), host defence polypeptides are evolutionarily conserved molecules, against which a widespread resistance has yet to emerge and cannot be developed easily. Resistance mechanisms against them exist but are not deemed systemic and are readily counteracted by relatively marginal alternations in peptide structure (*69, 70*). In this light, the pseudo-capsids demonstrated here may constitute a step change in the evolution of host defence effector molecules shifting the host-pathogen arms race in favor of more sustainable and adaptable antimicrobial treatments, serving an ultimate pursuit of reducing the spread of antimicrobial resistance.

REFERENCES

1. Podolsky, S. H. The evolving response to antibiotic resistance (1945-2018). *Palgrave Commun.* **4**, 124 (2018).
2. Tacconelli, E. et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis*. **18**, 318-327 (2018).
3. Laxminarayan, R., Matsoso, P., Pant, S., Brower, C., Røttingen, J. A., Klugman, K. and Davies, S. Access to effective antimicrobials: a worldwide challenge. *Lancet* **387**, 168-175 (2016).
4. Coates, A., Hu, Y., Bax, R. and Page, C. The future challenges facing the development of new antimicrobial drugs. *Nat. Rev. Drug Discov*. **1**, 895-910 (2002).
5. Haney, E. F., Straus, S. K. and Hancock R. E. W. Reassessing the host defense peptide landscape. *Front. Chem*. **7**, 43 (2019).
6. Hancock, R. E. and Sahl, H.-G. Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nat. Biotechnol*. **24**, 1551–1557 (2006).
7. Shai, Y. Mode of action of membrane active antimicrobial peptides. *Biopolymers* **66**, 236–248 (2002).
8. Wang, G. Human antimicrobial peptides and proteins. *Pharmaceuticals* **7**, 545-594 (2014).
9. Telang, S. Lactoferrin: a critical player in neonatal host defense. *Nutrients* **10**, 1228 (2018).
10. Hunter, H. N., Demcoe, A. R., Jenssen, H., Gutteberg, T. J. and Vogel, H. J. Human lactoferricin is partially folded in aqueous solution and is better stabilized in a membrane mimetic solvent. *Antimicrob. Agents Chemother*. **49**, 3387-3395 (2005).
11. Krauzewicz, N., Stokrová, J., Jenkins, C., Elliott, M., Higgins, C. F. and Griffin, B. E. Virus-like gene transfer into cells mediated by polyoma virus pseudocapsids. *Gene Ther*. **7**, 2122-2131 (2000).

12. Saugar, I., Luque, D., Oña, A., Rodríguez, J. F., Carrascosa, J. L., Trus, B. L. and Castón, J. R. Structural polymorphism of the major capsid protein of a double-stranded RNA virus: an amphipathic alpha helix as a molecular switch. *Structure* **13**, 1007-1017 (2005).
13. Bajaj, S. and Banerjee, M. In vitro assembly of polymorphic virus-like particles from the capsid protein of a nodavirus. *Virology* **496**, 106-115 (2016).
14. Uppalapati, M. et al. A potent D-protein antagonist of VEGF-A is nonimmunogenic, metabolically stable, and longer-circulating *in vivo*. *ACS Chem. Biol*. **11**, 1058-1065 (2016).
15. Arranz-Gibert, P., Ciudad, S., Seco, J., Garcia, J., Giralt, E. & Teixido, M. Immunosilencing peptides by stereochemical inversion and sequence reversal: retro-D-peptides. *Sci. Rep*. **8**, 6446 (2018).
16. Ryadnov, M. G., Degtyareva, O. V., Kashparov, I. A. and Mitin, Y. V. A new synthetic all-D-peptide with high bacterial and low mammalian cytotoxicity. *Peptides* **23**, 1869-1871 (2002).
17. King, T. P., Wade, D., Coscia, M. R., Mitchell, S., Kochoumian, L. and Merrifield, B. Structure-immunogenicity relationship of melittin, its transposed analogues, and D-melittin. *J. Immunol*. **153**, 1124-1131 (1994).
18. Bland, J. M., De Lucca, A. J., Jacks, T. J. and Vigo, C. B. All-D-cecropin B: synthesis, conformation, lipopolysaccharide binding, and antibacterial activity. *Mol. Cell Biochem*. **218**, 105-111 (2001).
19. García-Montoya, I. A., Cendón, T. S., Arévalo-Gallegos, S. and Rascón-Cruz, Q. Lactoferrin a multiple bioactive protein: an overview. *Biochim. Biophys. Acta* **1820**, 226-236 (2012).
20. Zhao, X. Y. and Hutchens, T. W. Proposed mechanisms for the involvement of lactoferrin in the hydrolysis of nucleic acids. *Adv. Exp. Med. Biol*. **357**, 271-278 (1994).
21. Sallmann, F. R., Baveye-Descamps, S., Pattus F., Salmon, V., Branza, N., Spik, G. and Legrand, D. Porins OmpC and PhoE of *Escherichia coli* as specific cell-surface targets of human lactoferrin. Binding characteristics and biological effects. *J. Biol. Chem*. **274**, 16107-16114 (1999).
22. Schibli, D. J., Hwang, P. M. and Vogel, H. J. The structure of the antimicrobial active center of lactoferricin B bound to sodium dodecyl sulfate micelles. *FEBS Lett*. **446**, 213-217 (1999).
23. Cochran, A. G., Skelton, N. J. and Starovasnik, M. A. Tryptophan zippers: stable, monomeric beta-hairpins. *Proc. Natl. Acad. Sci. USA*. **98**, 5578-5583 (2001).
24. Venkataram Prasad, B. V. and Schmid M. F. Principles of virus structural organization. *Adv. Exp. Med. Biol*. **726**, 17-47 (2012).
25. Kirchhausen, T. Clathrin. *Annu. Rev. Biochem*. **69**, 699-727 (2000).
26. Matsuura K. Synthetic approaches to construct viral capsid-like spherical nanomaterials. *Chem. Commun. (Camb)* **54**, 8944-8959 (2018).
27. Castelletto, V. et al. Structurally plastic peptide capsules for synthetic antimicrobial viruses. *Chem. Sci*. **7**, 1707–1711 (2016).
28. Fletcher, J. M. et al. Self-assembling cages from coiled-coil peptide modules. *Science* **340**, 595–599 (2013).
29. Mannige R. V. and Brooks C. L. 3rd. Periodic table of virus capsids: implications for natural selection and design. *PLoS One* **5**, e9423 (2010).
30. Schein, S. and Gayed, J. M. Fourth class of convex equilateral polyhedron with polyhedral symmetry related to fullerenes and viruses. *Proc. Natl. Acad. Sci. USA* **111**, 2920-2925 (2014).
31. Holowka, E. P., Sun, V. Z., Kamei, D. T. and Deming, T. J. Polyarginine segments in block copolypeptides drive both vesicular assembly and intracellular delivery. *Nat. Mater*. **6**, 52-57 (2007).
32. Oostenbrink, C., Villa, A., Mark, A. and Van Gunsteren, W. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6, *J. Comput. Chem*. **25**, 1656–1676 (2004).

33. Monticelli, L., Kandasamy, S. K., Periole, X., Larson, R. G., Tieleman, D. P. and Marrink, S. J. The MARTINI Coarse-Grained Force Field: Extension to Proteins, *J. Chem. Theory Comput.* **4**, 819–834 (2008).

34. Glasgow, J. and Tullman-Ercek, D. Production and applications of engineered viral capsids. *Appl. Microbiol. Biotechnol*. **98**, 5847–5858 (2014).

35. Maassen, S. J., van der Ham, A. M. and Cornelissen, J. L. M. Combining protein cages and polymers: from understanding self-assembly to functional materials. *ACS Macro Lett*. **5**, 987-994 (2016).

36. Noble, J. E. et al. A de novo virus-like topology for synthetic virions. *J. Am. Chem. Soc.*, **138**, 12202–12210 (2016).

37. Rakowska, P. D. et al. Nanoscale imaging reveals laterally expanding antimicrobial pores in lipid bilayers. *Proc. Natl. Acad. Sci. USA*, **110**, 8918–8923 (2013).

38. Pyne, A. et al. Engineering monolayer poration for rapid exfoliation of microbial membranes. *Chem. Sci*. **8**, 1105–1115 (2017).

39. Martínez, L., Andrade, R., Birgin, E. G. and Martínez, J. M. Packmol: a package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **30,** 2157-2164 (2009).

40. Wang, P., Robert, L., Pelletier, J. Dang, W. L., Taddei, F., Wright, A. and Jun, S. Robust growth of Escherichia coli. *Curr. Biol*. **20**, 1099-1103 (2010).

41. Pagliara, S., Chimerel, C., Langford, R., Aarts, D. G. and Keyser, U. F. Parallel sub-micrometre channels with different dimensions for laser scattering detection. *Lab Chip* **11**, 3365–3368 (2011).

42. Bamford, R. A., Smith, A., Metz, J., Glover, G., Titball, R. W. and Pagliara, S. Investigating the physiology of viable but non-culturable bacteria by microfluidics and time-lapse microscopy. *BMC Biol*. **15**, 121 (2017).

43. Balaban, N. Q. et al. Definitions and guidelines for research on antibiotic persistence. *Nat. Rev. Microbiol*. doi: 10.1038/s41579-019-0196-3 (2019)

44. Zhang, Y. Persisters, persistent infections and the Yin-Yang model. *Emerg. Microbes Infect.* **3**, e3 (2014).

45. Ayrapetyan, M., Williams, T. C. and Oliver, J. D. Bridging the gap between viable but non-culturable and antibiotic persistent bacteria. *Trends Microbiol*. **23**, 7–13 (2015).

46. Stapels, D. A. C., Hill, P. W. S, Westermann, A. J., Fisher, R. A., Thurston, T. L., Saliba, A. E., Blommestein, I., Vogel, J. and Helaine, S. *Salmonella* persisters undermine host immune defenses during antibiotic treatment. *Science* **362**, 1156-1160 (2018).

47. Schindler, P. R., and Teuber, M. Action of polymyxin B on bacterial membranes: morphological changes in the cytoplasm and in the outer membrane of *Salmonella typhimurium* and *Escherichia coli* B. *Antimicrob. Agents Chemother*. **8**, 95-104 (1975).

48. Edgar, R., Rokney, A., Feeney, M., Semsey, S., Kessel, M., Goldberg, M. B., Adhya, S. and Oppenheim, A. B. Bacteriophage infection is targeted to cellular poles. *Mol. Microbiol*. **68**, 1107-1116 (2008).

49. Hu, B., Margolin, W., Molineux, I. J. and Liu, J. Structural remodeling of bacteriophage T4 and host membranes during infection initiation. *Proc. Natl. Acad. Sci. USA* **112**, E4919-E4928 (2015).

50. Shaw, M., Zajiczek, L. and O'Holleran, K. High Speed Structured Illumination Microscopy in Optically Thick Samples. *Methods* **88**, 11−19 (2015).

51. O'Holleran, K. and Shaw, M. Optimized approaches for optical sectioning and resolution enhancement in 2D structured illumination Microscopy. *Biomed. Opt. Express* **5**, 2580−2590 (2014).

52. Shaw, M., Bella, A. and Ryadnov, M. G. CREIM: Coffee ring effect imaging model for monitoring protein self-assembly in situ. *J. Phys. Chem. Lett*. **8**, 4846-4851 (2017).

53. Desbois, A. P. and Coote, P. J. Wax moth larva (*Galleria mellonella*): an in vivo model for assessing the efficacy of antistaphylococcal agents. *J. Antimicrob Chemother,* **66**, 1785-1790, (2011).

54. Ba, X. *et al.* Old drugs to treat resistant bugs: methicillin-resistant *Staphylococcus aureus* isolates with mecC are susceptible to a combination of penicillin and clavulanic acid. *Antimicrob. Agents Chemother.* **59**, 7396-7404, (2015).

55. Gibreel, T. M. and Upton, M. Synthetic epidermicin NI01 can protect *Galleria mellonella* larvae from infection with *Staphylococcus aureus*. *J. Antimicrob. Chemother*. **68**, 2269-2273 (2018).

56. Takemura-Uchiyama, I., Uchiyama, J., Kato, S., Inoue, T., Ujihara, T., Ohara, N., Daibata, M. and Matsuzaki, S. Evaluating efficacy of bacteriophage therapy against *Staphylococcus aureus* infections using a silkworm larval infection model. *FEMS Microbiol. Lett*. **347**, 52-60 (2013).

57. De Santis, E, et al. Antimicrobial peptide capsids of de novo design. *Nat Commun*. **8**, 2263 (2017)

58. Matsuura, K. Construction of spherical virus-inspired peptide nanoassemblies. *Polym. J.* **44**, 469–474 (2012).

59. Tarasov, S. G. et al. Structural plasticity of a transmembrane peptide allows self-assembly into biologically active nanoparticles. *Proc. Natl. Acad. Sci. USA* **108**, 9798–9803 (2011).

60. Ryadnov, M. G. A self-assembling peptide polynanoreactor. *Angew. Chem. Int. Ed.* **46**, 969–972 (2007).

61. Papapostolou, D., Smith, A. M., Atkins, E. D., Oliver, S. J., Ryadnov, M. G., Serpell, L. C. and Woolfson, D. N. Engineering nanoscale order into a designed fiber. *Proc. Natl. Acad. Sci USA* **104**, 10853-10858 (2007).

62. Matsuura, K., Murasato, K. and Kimizuka, N. Artificial peptide nanospheres self-assembled from three-way junctions of β-sheet-forming peptides. *J. Am. Chem. Soc.* **127**, 10148–10149 (2005).

63. Brauner, A., Fridman, O., Gefen, O. and Balaban, N. Q. Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nat. Rev. Microbiol*. **14**, 320-330 (2016).

64. Neuhaus, F. C. and Baddiley, J. A continuum of charge: structures and functions of D-alanyl-teichoic acids in Gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* **67**, 686–723 (2003).

65. Needham, B. D. and Trent, M. S. Fortifying the barrier: the impact of lipid A remodeling on bacterial pathogenesis. *Nat. Rev. Microbiol.* **11**, 467–481 (2013).

66. Joo, H.-S., Fu, C.-I. and Otto, M. Bacterial strategies of resistance to antimicrobial peptides. *Phil. Trans. R. Soc. B* **371**, 20150292 (2016).

67. Ryan, L. *et al*. Anti-antimicrobial peptides: folding-mediated host defense antagonists. *J. Biol. Chem.* **288**, 20162–20172 (2013).

68. Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature* **415**, 389–395 (2002).

69. Lazar, V. et al. Antibiotic-resistant bacteria show widespread collateral sensitivity to antimicrobial peptides *Nat. Microbiol*. **3**, 718-731 (2018).

70. Pfeil, M. P. et al. Tuneable poration: host defense peptides as sequence probes for antimicrobial mechanisms. *Sci. Rep*. **8**, 14926 (2018).

Supplementary information contains experimental notes together with additional data, table and figures

Acknowledgements

We acknowledge funding from the UK's Department for Business, Energy and Industrial Strategy. UL and SP were supported through an MRC Proximity to Discovery EXCITEME2 grant (MCPC17189).

Author contributions

All authors designed and performed the experiments. M.G.R wrote the manuscript. All authors analysed data and contributed to the editing of the manuscript.

Additional Information

The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.G.R.

Supporting Information

Engineering persistent antimicrobial pseudo-capsids

Ibolya E Kepiro, Irene Marzuoli, Katharine Hammond, Xiaoliang Ba, Helen Lewis, Michael Shaw, Smita B Gunnoo, Emiliana De Santis, Urszula Łapińska, Stefano Pagliara, Mark Holmes, Christian D Lorenz, Bart W Hoogenboom, Franca Fraternali and Maxim G Ryadnov

Materials and Methods

Peptide synthesis and purification: L- and D-triskelions – (RRWTWE)-βA-K(RRWTWE)-K(RRWTWE)-am – were assembled on a Liberty microwave peptide synthesizer (CEM) using standard Fmoc/^{t}Bu solid-phase protocols with HBTU/DIPEA as coupling reagents on a Rink amide MBHA resin. Fmoc-K(Mtt)-OH was used to enable orthogonal conjugation via a tri-functional dendritic hub – βAKK-am. Carboxyfluorescein-labelled triskelions were made on resin by coupling carboxyfluorescein to the N-terminus of βA-RRWTWE-βAKK, with fully protected side chains, on the resin. The remaining two arms were then assembled on resin following Mtt removal. Magainin 2 and cecropin B were synthesized as peptide amides on a Rink amide MBHA resin and LL37 was assembled as a peptide acid on a Fmoc-S(^{t}Bu)-Wang resin. After post-synthesis cleavage and deprotection (95% TFA, 2.5% TIS, 2.5% water) all peptides were purified by semi-preparative reversed-phase high-performance liquid chromatography (RP-HPLC). Peptide identities were confirmed by analytical RP-HPLC and MALDI-ToF.

MS $[M + H]^+$: cecropin B – *m/z* 3834.6 (calc.), 3836.0 (found); LL37 – *m/z* 4493.2 (calc.), 4495.8 (found); magainin 2 – *m/z* 2465.9 (calc.), 2467.0 (found); L-triskelion – m/z 3090.5 (calc.), 3092.3 (found); D-triskelion – m/z 3090.5 (calc.), 3092.8 (found); carboxyfluorescein-L-triskelion – m/z 3518.9 (calc.), 3519.7 (found).

Analytical and semi-preparative RP-HPLC was performed on a JASCO HPLC system (PU-980; Tokyo, Japan), using a Vydac C18 analytical and semi-preparative (both 5 μm) columns. Both analytical and semi-preparative runs used a 10-70% B gradient over 30 min at 1 mL/min and 4.5 mL/min, respectively, with detection at 280 and 214 nm (buffer A, 5% and buffer B, 95% aqueous CH_3CN, 0.1% TFA).

Ψ-capsid assembly: triskelions were assembled overnight at the concentrations stated in the text (100-400 μM) in filtered (0.22 μm), 10 mM MOPS, phosphate or PBS buffers, pH 7.4, at room temperature.

Circular dichroism spectroscopy: CD spectra were recorded on a JASCO J-810 spectropolarimeter fitted with a Peltier temperature controller. The measurements were taken in ellipticities in mdeg and converted to molar ellipticities ($[\theta]$, deg cm^2 $dmol^{-1}$ res^{-1}) by normalizing for the concentration of peptide bonds. The data was collected with a 1 nm step, 1 sec collection time per step, and is presented as the average of 4 scans.

Fourier transform infra-red spectroscopy: FT-IR spectra were recorded using a Tensor-27 series FTIR spectrometer equipped with a BioATR II unit (Bruker Optics), as the sampling platform, and a photovoltaic mercury cadmium telluride (MCT) detector and a Bruker Optics workstation equipped with OPUS software. Low-volume (20 μL) capsid samples (100 μM) were placed in a circular sampling area of radius 2 mm with a path length of 6 μm. FTIR spectra was recorded with resolution 4 cm^{-1}, scanner velocity 20 kHz, 256 scans, phase resolution 32 and zero filling factor 4. Spectra deconvolutions were performed by Gaussian peak fitting using the proprietary software.

Molecular dynamics simulations: MD simulations were performed using the GROMACS 2016 software with the GROMOS96 53A6 force field (atomistic) and the MARTINI force field (coarse grain) (*32*, *33*). A SPC water model was used for atomistic simulations and standard MARTINI water for coarse grain simulations (*33*). In all simulations chloride and sodium ions were added up to a 150 mM concentration. Additional ions were placed if required for charge neutralization. The initial Ψ-capsid configuration was constructed geometrically using the editconf tool from GROMACS. The DLPC/DLPG (3:1) membrane was constructed with the PACKMOL software with dimensions of 12 × 12 nm (atomistic) or 30 × 30 nm (coarse grain) (*39*). Periodic boundary conditions were imposed, setting the minimal distances between the protein and the box boundaries to 3 nm. All simulations are run in the NPT ensemble maintaining the temperature at 303K with the velocity rescale thermostat, and the pressure at 1 bar using the Parrinello-Rahman barostat, with either an isotropic or semi-isotropic coupling; simulations without and with membranes, respectively. For atomistic simulations of the Ψ-capsid in solution, multiple restrained equilibration runs (65 ns overall) precede the 100 ns production, which were run in three replicates. Similarly, for coarse grain simulations a 10 ns equilibration precedes the 1 μs production, which were run in three replicates. For simulations including membranes, the lipid patches were equilibrated for 500 ns (1 μs for coarse grain) before placing the peptide close in contact with the lipid surface (atomistic) or at the minimal distance of 1 nm (coarse grain). Production was then carried out for 500 ns, and 10 μs for coarse grain. Atomistic simulations with membranes were run with an electric

field of 130 mV/nm in the direction perpendicular to the membrane, with the higher potential region on the side of the bilayer where the peptide rests.

Transmission electron microscopy: micrographs of the Ψ-capsids were recorded using a JEOL 1010 transmission electron microscope equipped with an Orius SC1000 (Gatan Inc.) CCD camera, operated at 80 keV. Droplets of capsid solution were placed on glow discharged Cu finder grids coated with carbon film (mesh 200), stained with uranyl acetate (aq. 2%, w/v) for a few seconds and buffer excess was removed by blotting paper. Micrographs of *E. coli* were recorded using FEI Tecnai T12 transmission electron microscope equipped with a Morada G2 (Olympus Inc.) camera, operated at 80 keV. Samples were fixed with glutaraldehyde (aq. 5%, w/v), treated with a post-fixation agent osmium tetroxide (1%, w/v, 100 mM PIPES, pH 7.2) and stained with uranyl acetate (aq. 2%, w/v). The resulting samples were embedded in a Spurr resin, and were left to solidify at 60 ºC, over 24 hrs. The resin was then microtomed to ultrathin sections that were placed on Formvar carbon coated grids (mesh 200), post-stained using lead citrate (aq. 5%, w/v) and imaged.

Preparation of unilamellar phospholipid vesicles for AFM imaging: 1,2-dilauroyl-sn-glycero-3-phosphocholine (DLPC) with 1,2-dilauroyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DLPG) lipids used for vesicle construction were from Avanti Polar Lipids (Alabaster, USA). DLPC were used as mammalian model membranes, and DLPC/DLPG (3:1, molar ratios) were used as bacterial model membranes. The lipids were weighted up, dissolved in chloroform-methanol (2:1, vol/vol), dried under a nitrogen stream and then under vacuum to form a thin film. The film was hydrated in 10 mM phosphate buffer (pH 7.2) with shaking (1 h, 220 rpm) and bath sonicated for 30 min. The obtained suspension was extruded using a hand-held extruder (Avanti Polar lipids) (twenty nine times, polycarbonate filter, 0.05 µm) to give a clear solution of small unilamellar vesicles, whose sizes (50 nm) were confirmed by photon correlation spectroscopy.

In-liquid atomic force microscopy on supported lipid bilayers: supported lipid bilayers were formed using a vesicle fusion method as described elsewhere (*38*). Freshly prepared vesicles (1.5 µL, 3 mg/mL) were added to cleaved mica a pre-hydrated with buffer (70 µL, 120 mM NaCl, 20 mM MOPS, 20 mM $MgCl_2$). Following adsorption and incubation for 45 min, the samples were washed to remove unfused vesicles and resulting SLBs were checked to confirm they were defect free. Peptides were then introduced into a 100-µL fluid cell (Bruker AXS, USA) where they were diluted to the final concentrations of 0.1-0.8 µM. All imaging was performed using Peak Force Tapping™ mode on a Multimode 8 AFM system (Bruker AXS, USA) using MSNL-E cantilevers (Bruker AFM probes, USA). Images were taken at PeakForce frequency of 2 kHz, PeakForce amplitude of 10 nm and PeakForce set-point of 10-30 mV (<100 pN). Images were processed using Nanoscope analysis software (Bruker AXS, USA) or Gwyddion (http://gwyddion.net) for first order line-by-line background subtraction (flattening) and first-order plane fitting.

Photon correlation spectroscopy. Prepared phospholipid vesicles for AFM imaging were re-suspended to a final concentration of 1 mg/mL and were analysed on a Zetasizer Nano (ZEN3600; Malvern Instruments). Dynamic light scattering batch measurements were carried out in a low volume disposable cuvette at 25 °C. Hydrodynamic radii were obtained through the fitting of autocorrelation data using the manufacturer's Dispersion Technology Software (version 5.10).

Minimum inhibitory concentrations assay: minimum inhibitory concentrations were determined by broth microdilution on *P. aeruginosa*, *E. coli*, *S. aureus*, *M. luteus*, *B. subtilis*, *S. typhimurium* and *K. pneumoniae* according to the Clinical and Laboratory Standards Institute. Typically, 100 µL of 0.5–1 × 10^6 CFU per ml of each bacterium in Mueller Hinton media broth (Oxoid) were incubated in 96-well microtiter plates with 100 µL of serial two-fold dilutions of the capsids (from 100 to 0 µM) at 37 °C on a 3D orbital shaker. The absorbance was measured after capsid addition at 600 nm using a SpectraMax i3x Multi-Mode Microplate Reader (Molecular Devices). MICs were defined as the lowest peptide concentration that inhibited visible bacterial growth after 24 h at 37 °C. All tests were done at least in triplicate and results are summarized in Table S1. The values that are given without standard deviations are those for which no variations were found within triplicates.

Hemolysis assay: hemolysis was determined using human erythrocytes sourced commercially from Cambridge Bioscience Ltd. and used within two days. 10% (vol/vol) suspensions of human erythrocytes were incubated with the capsids. The cells were rinsed four times in 10 mM phosphate buffer saline (PBS, Gibco™), pH 7.2, by repeated centrifugation and re-suspension (3 min at 3000 × g). The cells were then incubated at room temperature for 1 h in either deionized water (fully hemolysed control), PBS, or with capsid in PBS. After centrifugation at 10,000 × g for 5 min, the supernatant was separated from the pellet, and the absorbance was measured at 550 nm using a SpectraMax i3x Multi-Mode Microplate Reader (Molecular Devices). Absorbance of the suspension treated with deionized water defined complete hemolysis. All tests were done in triplicate and results are shown in Table S1. The values given correspond to concentrations needed to kill a half of the sample population (50% lysis of erythrocytes) and are expressed as median lethal concentrations, LC_{50}.

Bacterial culture for single-cell antimicrobial kinetics measurements: lysogeny broth (LB) (10 g/L tryptone, 5 g/L yeast extract, and 0.5 g/L NaCl, Formedium) and LB agar plates (LB with 15 g/L agar) were used for planktonic and plate bacteria growth, respectively. *E. coli* BW25113 was purchased from Dharmacon (GE Healthcare). Single colonies were picked from a streak plate and

were incubated at 37 °C, over 17 hours in fresh LB medium (100 mL) with shaking (200 rpm). The prepared culture was span down (10 min, 3000 g and 20 ºC). The supernatant was filtered twice (Medical Millex-GS Filter, 0.22 µm, Millipore Corp.) and used to re-suspend the bacteria pellet to an OD_{595} of 50. The resulting highly concentrated bacterial suspension was used in high throughput single-cell kinetic measurements, together with M9 minimal medium (7 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 1 g/L NH_4Cl, 0.5 g/L NaCl, 1 mM thiamine hydrochloride) (Sigma Aldrich), ampicillin (Sigma Aldrich) and propidium iodide (PI) (Thermo Fisher Scientific), as described below.

High throughput single-cell kinetic measurements: antimicrobial efficacy with a single-cell resolution was measured using a multi-channel microfluidics device as reported elsewhere (*40, 41*). The device used was a polydimethylsiloxane (PDMS, Dow Corning) replica of the original mold kindly provided by Suckjoon Jun (*40*). The device is an array of dead-end microfluidic channels 1.5 µm in diameter and a 25 µm in length. Each channel can host 1-4 bacterial cells in a single file. The channels are connected to a main microchamber (25x100 µm) that is continuously supplied with antimicrobial agents, fresh LB or propidium iodide (see below). The device was permanently attached to a glass coverslip and functionalized with a solution of bovine serum albumin (50 mg/mL). 2 µl of the highly concentrated bacterial suspension, prepared as described above, was injected in the device and individual bacteria were allowed to diffuse into the lateral channels for 30 min. The device was then completed with fluorinated ethylene propylene inlet and outlet tubing (1/32"×0.008") connected to a flow-rate measuring system (Flow Unit S, Fluigent, France) with the applied pressure controlled by a computerized pressure-based flow control system (MFCS-4C, Fluigent). The completed device was mounted on an inverted microscope (IX73 Olympus, Tokyo, Japan) equipped with a 60×, 1.2 N.A. objective (UPLSAPO60XW, Olympus) and a sCMOS camera (Zyla 4.2, Andor, Belfast, UK) used to acquire bright-field images (exposure time 0.03s) of thousand channels per hour for the duration of each experiment (24 hrs). After acquiring the first set of images, the microfluidic environment was changed by flowing M9:LB (9:1 v:v) containing antimicrobial agents (D/Ψ-capsids or ampicillin) at 25×MIC at 100 µL/h over three hours and then by LB for a further 21 hrs. At 24 h PI (30 µM) was delivered into the microfluidic device for 10 min to identify dead cells with compromised cell membranes. The device allowed to track each bacterial cell and its progeny throughout the assay, with the images obtained in each channel loaded to ImageJ.

Time-lapse structured illumination microscopy of *E. coli* and *S. Aureus*: bacteria (*E. coli* ATCC 15597 and *S. aureus* ATCC 6538) were inoculated into Mueller Hinton broth, grown overnight at 37 °C, 150 rpm, re-inoculated and grown to OD_{600} of 0.5-0.8. The cells were then pelleted by centrifugation (4000 rpm, 10 mins), and after removing the supernatant, were re-suspended in 1 mL PBS (10 mM) and transferred to an Eppendorf tube. FM4-64 (1 mg/mL, 5 µL) was added to each tube followed by rotation on a rotating disk for 5 mins. The cells were then centrifuged (7000 rpm, 3 mins), the supernatant was removed, and the cell pellet was re-suspended in PBS (1 mL). Bacteria was then diluted to OD_{600} of 0.16 and 0.08 for *E. coli* and *S. aureus*, respectively. 100 µL of each suspension was added to different wells of a chambered microscope coverslip (Ibidi µ-Slide 8 Well) and left for 30 min to allow cells to settle onto the substrate. Immediately prior to each imaging experiment a 100-µL solution of Ψ-capsids assembled from caroxyfluorescein-labelled and unlabelled triskelions at 1:100 molar ratios were added to the well to a final concentration of 50 µM (total peptide). For control datasets, 100 µL of PBS was added prior to imaging. Time lapse imaging was performed at 5-min intervals using a custom-built structured illumination microscope (*50*). Briefly, sinusoidal excitation patterns were generated by projecting a spatially filtered image of a spatial light modulator (SLM), configured to display a series of binary phase gratings, into the focal plane of the microscope objective lens (UPLSAPO 60x/1.3, Olympus). Sample images were acquired using a scientific CMOS camera (ORCA-Flash4.0, Hamamatsu Photonics), with the global exposure period of the camera's rolling shutter synchronised to the pattern displayed on the SLM. At each time point FM4-64 and carboxyfluorescein-labelled capsids were imaged sequentially, using excitation at 561 nm with a 655 nm long pass emission filter and excitation at 488 nm with a 525/50 bandpass emission filter respectively. Images were reconstructed as described elsewhere (*51*) and lateral drift and chromatic offsets in the reconstructed images were corrected using ImageJ. The FM4-64 colour channel was then corrected for photobleaching using an exponential fit to the image intensity within a suitable region of interest. Manual background subtraction and colour balancing was performed to optimise image contrast.

In vivo MRSA clearance assay: *Galleria mellonella* larvae assay was performed as described elsewhere (*53*, *54*). Wax moth larvae were purchased in bulk from Livefood UK, stored at 4°C upon arrival and kept at 37°C during the course of the assay. Ψ-capsid samples were prepared from stock solutions in Milli-Q water to working concentrations in sterile phosphate buffered saline (PBS). Typically, a single bacterial colony was picked to inoculate 5 mL of tryptic soy broth (TSB), and the resulting culture was left to grow overnight (~16 hours) at 37°C with shaking (200 rpm). The culture was then diluted (100x) into a fresh TSB (5 mL) and incubated over 4 hours at 37°C with shaking (200 rpm). After centrifugation (2500g) for 10 min, pellets were obtained and re-suspended in sterile PBS to an OD_{595} of 0.2, giving ~10^8 CFU/mL.) The re-suspended culture (10 µL aliquots; ~1.2×10^6 CFU) was administered to larvae (seven groups; n=10 in each group) behind the rear thoracic segments using a Tridak Stepper Pipette Dispenser (Dymax, UK). The infected larvae were then treated by injection with 10 µL of L/Ψ-capsid (12.5 µM), D/Ψ-capsid, (12.5 µM),

magainin 2 (12.5 μM), cecropin B (12.5 μM), LL37 (12.5 μM), vancomycin (40 μM) or PBS at stated time points after inoculation. The toxicity of Ψ-capsid was assessed using two groups of larvae (n=10 in each group). These were mock-infected with PBS and treated with 10 μL of L or D/Ψ-capsid (25 μM) at stated time points after inoculation. For the delayed treatment assay, three groups of infected larvae (~10^6 CFU) were treated with 10 μL of L/Ψ-capsid (12.5 μM), vancomycin (40 μM) or PBS at stated time points after inoculation. All experiments were carried out as blind studies, and the treatment identities were not revealed until each experiment was completed. Larvae were considered dead when they did not respond to touch to the head. Survival curves were generated and analysed using GraphPad Prism 6 software.

Table, Figures and Movies

Table S1. Biological activities of Ψ-capsids in comparison with other antimicrobial agents.

Cell	Antimicrobial agent					
	Ψ-capsids		ampicillin	melittin	polymyxin B	cecropin B
	L-form	D-form				
	Minimum Inhibitory Concentration, μM					
E. coli	7.8 ± 2.7	6.6 ± 3.9	>25	2.3 ± 1.1	1.3 ± 0.3	<1
S. aureus	8.3 ± 2.9	17.2 ± 8.1	<1	<1	25	50
P. aeruginosa	9.4 ± 3.1	>25	>25	9.3 ± 3.1	1.5	25
S. typhimurium	12.5	9.4 ± 5.4	8.3 ± 2.9	3.1	1.5	<1
K. pneumoniae	>25	8.2 ± 4.6	12.5	3.1	2.6 ± 0.7	1.5
B. subtilis	2.3 ± 0.8	5.2 ± 1.5	25	3.1	3.1	>50
M. luteus	1.2 ± 0.4	2.3 ± 0.8	1 ± 0.4	<1	2 ± 0.7	<1
	LC_{50},[a] μM					
Human erythrocytes	>250	>250	n.d.	<10	>250	>250

[a]concentration required to achieve 50% cell death compared to untreated cells

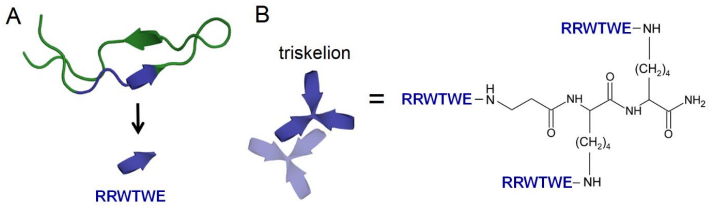


Figure S1. Triskelion design. (a) N-terminal domain of lactoferrin (PDB entry 1LFC rendered by PyMol). The core sequence, RRWQWR, in the domain is highlighted in blue, and converted to a self-complementary β-strand, RRWTWE. (B) Three copies of RRWTWE are conjugated into a β-strand triskelion. Triskelions interact via β-strand arms. For clarity only two triskelions forming a β-sheet are shown and highlighted in different colours. (C) The chemical structure of the triskelal RRWTWE conjugate. Note: the schematics are the same for the D-form of the triskelion, with all amino acids being D-epimers.

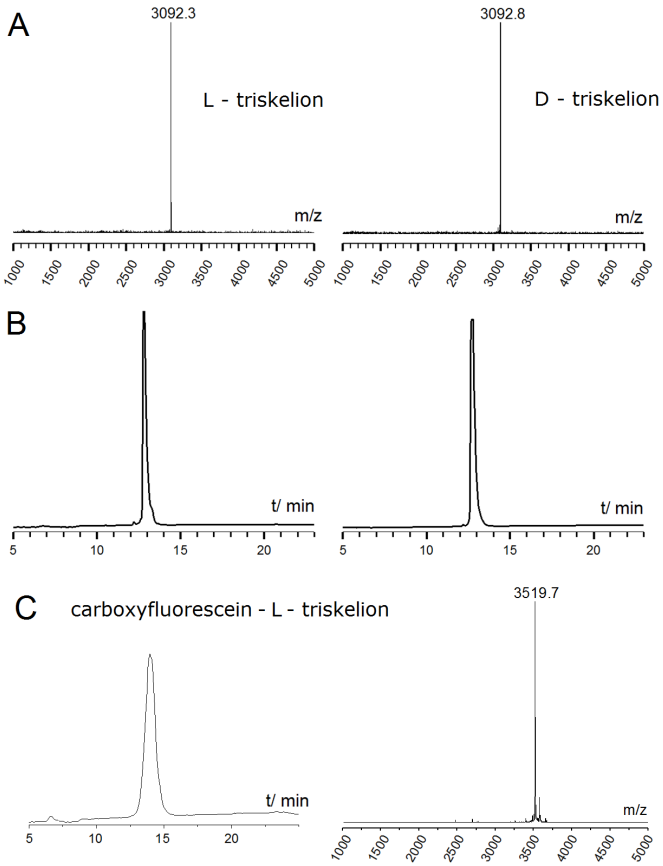


Figure S2. Triskelion synthesis. MALDI-ToF spectra (A) and RP-HPLC profiles (B) for purified L- and D-triskelions. (C) RP-HPLC (left) and MALDI-ToF (right) traces for the purified L-triskelion modified with carboxyfluorescein at the N-terminus of one arm.

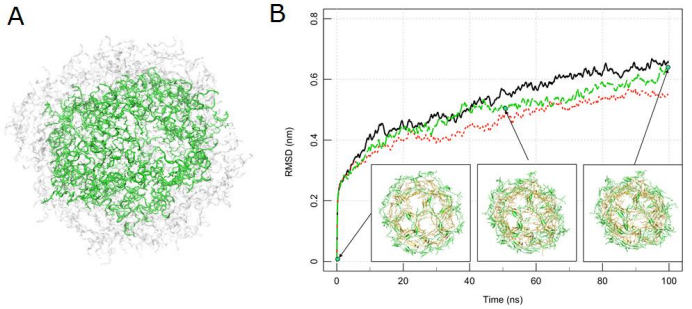


Figure S3. Pseudo-capsid shells. (A) Final snapshot (1 μs) of a coarse grain MD simulations of a β-sheet single-walled shell templated on a truncated icosahedron (green, peptide backbone only). For comparison, a final configuration for double-walled shell is shown in grey (peptide backbone only). (B) Root mean square deviations (RMSD) of double-walled shells versus their initial configurations as a function of time. Initial, mid-time and final configurations are shown in insets. See also Movie S1.

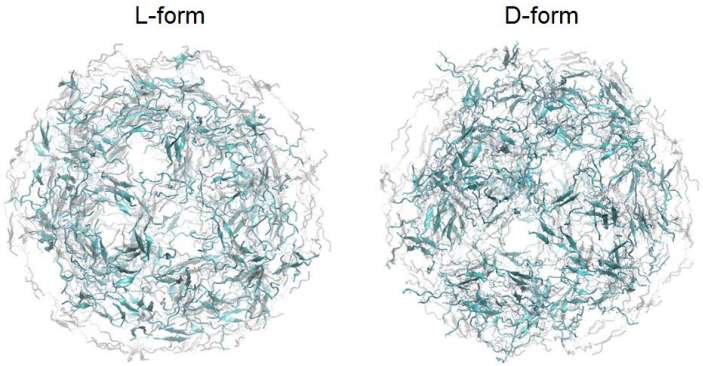


Figure S4. Double-walled pseudo-capsid shell. Final configurations of L- and D-forms of the triskelion assembly templated on a truncated isosahedron after 100 ns of atomistic simulations. Initial configurations (peptide backbone only) are shown in grey.

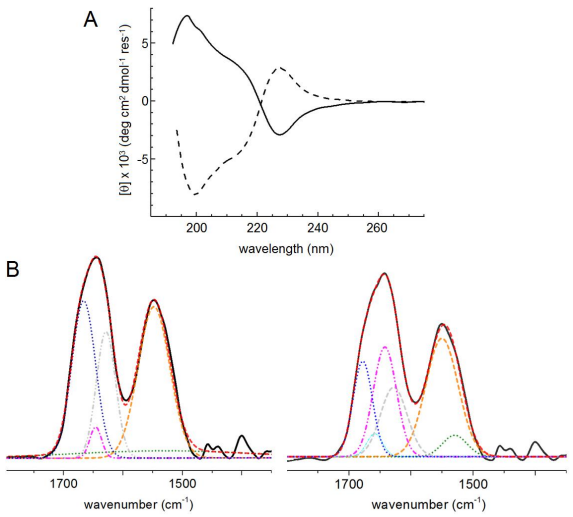


Figure S5. Pseudo-capsid folding. (A) CD spectra for L/Ψ-capsids (dashed line) and D/Ψ-capsids (solid line). (B) FT-IR spectra for L/Ψ-capsids (right) and D/Ψ-capsids (left), including raw spectra (black line), cumulative fit (dashed red line), deconvoluted amide spectra (other coloured lines). Folding conditions: 100 μM peptide, pH 7.4, 10 mM MOPS, 20˚C.

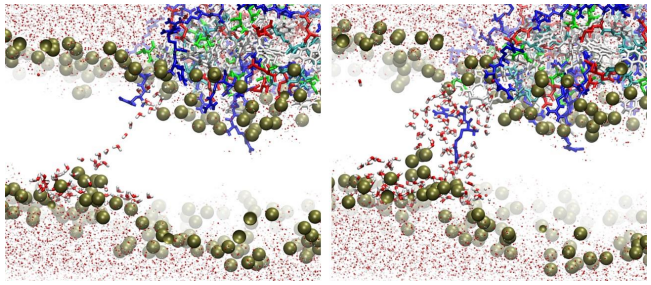


Figure S6. A rudimentary pore. Snapshots taken at 62 ns (left) and 72 ns (right) of 100-ns atomistic simulations of a rudimentary pore in phospholipid bilayers. Key: for clarity, only phosphate groups (golden beads) are shown for the bilayer. See also Movies S2 and S3.

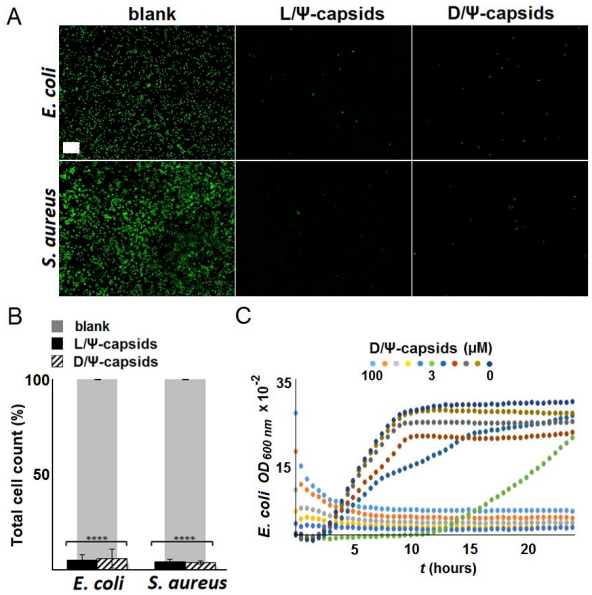


Figure S7. Antimicrobial kinetics of Ψ-capsids. (A) Fluorescence micrographs of bacterial cells after 16-hour incubations without (blank) and with Ψ-capsids. The scale bar is 20 μm. (B) Total cell counts for samples after 16-hour incubations without (blank), taken as 100%, and with Ψ-capsids. Total cell counts for samples treated with the capsids were significantly lower than those for un-treated cells (p <0.001). (C) Representative bacteria growth curves (*E. coli*) recorded for serial peptide dilutions (100 μM to 0 μM) as a function of time. Individual data points are collected at 30 min intervals.

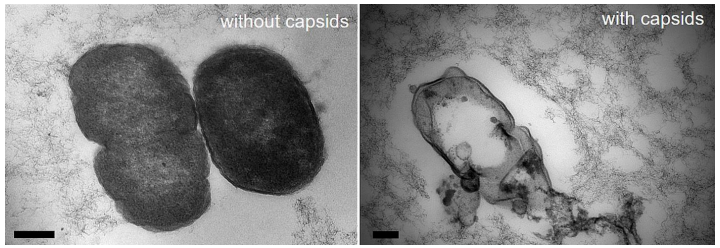


Figure S8. Single-cell antimicrobial kinetics of Ψ-capsids. Electron micrographs of microtomed *E. coli* cells without and after treatment with D/Ψ-capsids. The scale bars are 200 nm.

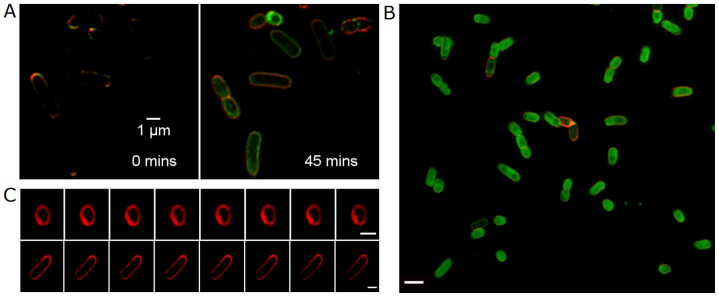


Figure S9. Time-kill kinetics of pseudo-capsids in vitro. (A) SIM images of *E. coli* cells immediately after the addition of L/Ψ-capsids and after 45 min of the treatment. The capsids are labelled with carboxyfluorescein (green). Cell membranes are stained with FM4-64FX (red). (B) A larger area of *E. coli* cells imaged at 45 min of the treatment. (C) Control SIM images recorded at 5-min intervals for *S. aureus* (upper) and *E. coli* (lower) without the capsids. Scale bars are 1 µm.

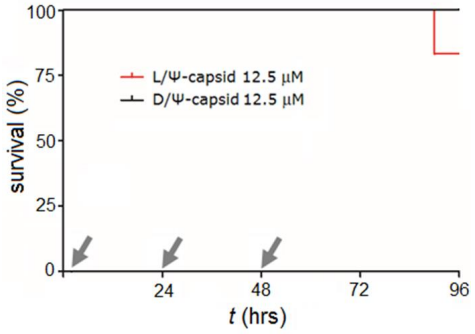


Figure S10. Time-kill kinetics of pseudo-capsids in vivo. Survival of *G. mellonella* larvae treated with Ψ-capsids. Inoculations were done straight after the initiation of infection (first 2 hrs) with two subsequent treatments at 24 and 48 hrs.

A.2 Supplementary material

Supplementary material can be found at `https://github.com/imarzuoli/Thesis_marzuoli` and includes Supplementary Movies of:

- `SI_M1_GROMOS_buckyball.mp4`
  GROMOS atomistic simulation of the buckyball bilayer - 100 ns;
- `SI_M2_SIRAH_buckyball.mpg` and `SI_M3_PolarMARTINI_buckyball.mpg`
  SIRAH and Polar MARTINI coarse-grained simulations of the buckyball bilayer and monolayer - 1 μs;
- `SI_M4_GROMOS_capzip_bacterial_E130_poration.mp4`
  atomistic simulation of poration of a model bacterial model membrane under the combined action of the pentagonal capzip subunit and an external electric field of 130 mV/nm - 60 ns;
- `SI_M5_MARTINI_capzip_bacterial.mp4` and `SI_M6_MARTINI_capzip_mammalian.mpg`
  standard MARTINI simulation of a buckyball bilayer approaching a bacterial and mammalian model membrane - 10 μs;
- `SI_M7_PolarMARTINI_capzip_bacterial_E40_poration.mpg`
  Polar MARTINI simulation of a buckyball bilayer porating a model bacterial membrane under the effect of a 40 mV/nm external electric field - 200 ns.

A.3 Software developed

The simulations and analysis of capzip systems prompted the development of tools to facilitate the task. These are meant to be used in conjunction with standard analysis tools provided by GROMACS [Berendsen et al., 1995; Abraham et al., 2015, 2018] and MDAnalysis [Michaud-Agrawal et al., 2011; Gowers et al., 2016]. The scripts and packages developed can be found at `https://github.com/imarzuoli/MDtools`, which includes (among others):

- a python module to manipulate and prepare multi-branched peptide topology files in the GROMACS format for GROMOS simulations (`MD_mutate_and_manipulate.py`);
- python functions to perform the analysis of contacts as in Section 3.4 of Chapter 3 (`contacts_analysis.py`);
- functions to perform the hydration analysis as in Section 3.4 of Chapter 4 (`an_min_hist.py`).

Bibliography

Abi Mansour, A. and Ortoleva, P. J. (2014). Multiscale Factorization Method for Simulating Mesoscopic Systems with Atomic Precision. *Journal of Chemical Theory and Computation*, 10(2):518–523.

Abouzeed, Y. M., Baucheron, S., and Cloeckaert, A. (2008). ramR mutations involved in efflux-mediated multidrug resistance in Salmonella enterica serovar Typhimurium. *Antimicrobial Agents and Chemotherapy*, 52(7):2428–34.

Abraham, E. P. and Chain, E. (1988). An enzyme from bacteria able to destroy penicillin. 1940. *Reviews of Infectious Diseases*, 10(4):677–8.

Abraham, M. J., Murtola, T., Schulz, R., et al. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19–25.

Abraham, M. J., van der Spoel, D., Lindahl, E., et al. (2018). *GROMACS User Manual version 2016*. www.gromacs.org.

Aguilera, O., Ostolaza, H., Quirós, L., et al. (1999). Permeabilizing action of an antimicrobial lactoferricin-derived peptide on bacterial and artificial membranes. *FEBS Letters*, 462(3):273–277.

Alom, M. Z., Taha, T. M., Yakopcic, C., et al. (2019). A State-of-the-Art Survey on Deep Learning Theory and Architectures. *Electronics*, 8(3):292.

Anézo, C., de Vries, A. H., Höltje, H.-D., et al. (2003). Methodological Issues in Lipid Bilayer Simulations. *The Journal of Physical Chemistry B*, 107(35):9424–9433.

Anguela, X. M. and High, K. A. (2019). Entering the Modern Era of Gene Therapy. *Annual Review of Medicine*, 70(1):273–288.

Anjana, R., Vaishnavi, M. K., Sherlin, D., et al. (2012). Aromatic-aromatic interactions in structures of proteins and protein-DNA complexes: a study based on orientation and distance. *Bioinformation*, 8(24):1220–4.

Arnold, R., Cole, M., and McGhee, J. R. (1977). A bactericidal effect for human lactoferrin. *Science*, 197(4300):263–265.

Arnold, R. R., Brewer, M., and Gauthier, J. J. (1980). Bactericidal activity of human lactoferrin: sensitivity of a variety of microorganisms. *Infection and Immunity*, 28(3):893–8.

Arranz-Gibert, P., Ciudad, S., Seco, J., et al. (2018). Immunosilencing peptides by stereochemical inversion and sequence reversal: retro-D-peptides. *Scientific Reports*, 8(1):6446.

Arseneault, M., Bédard, S., Boulet-Audet, M., et al. (2010). Study of the Interaction of Lactoferricin B with Phospholipid Monolayers and Bilayers. *Langmuir*, 26(5):3468–3478.

Asmar, A. T. and Collet, J.-F. (2018). Lpp, the Braun lipoprotein, turns 50 - major achievements and remaining issues. *FEMS Microbiology Letters*, 365(18).

Aurenhammer, F. and Franz (1991). Voronoi diagrams - a survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23(3):345–405.

Balatti, G. E., Ambroggio, E. E., Fidelio, G. D., et al. (2017). Differential Interaction of Antimicrobial Peptides with Lipid Structures Studied by Coarse-Grained Molecular Dynamics Simulations. *Molecules*, 22(10).

Band, V. and Weiss, D. (2014). Mechanisms of Antimicrobial Peptide Resistance in Gram-Negative Bacteria. *Antibiotics*, 4(1):18–41.

Barducci, A., Bonomi, M., and Parrinello, M. (2010). Linking Well-Tempered Metadynamics Simulations with Experiments. *Biophysical Journal*, 98(9):L44–L46.

Barducci, A., Bonomi, M., and Parrinello, M. (2011). Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(5):826–843.

Barrera, E. E., Machado, M. R., and Pantano, S. (2019). Fat SIRAH: Coarse-grained phospholipids to explore membrane-protein dynamics. *Journal of Chemical Theory and Computation*. DOI: 10.1021/acs.jctc.9b00435.

Bellamy, W., Takase, M., Yamauchi, K., et al. (1992). Identification of the bactericidal domain of lactoferrin. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1121(1-2):130–136.

Bennett, W. F. D., Hong, C. K., Wang, Y., et al. (2016). Antimicrobial Peptide Simulations and the Influence of Force Field on the Free Energy for Pore Formation in Lipid Bilayers. *Journal of Chemical Theory and Computation*, 12(9):4524–4533.

Berendsen, H., Postma, J., Van Gunsteren, W., et al. (1981). Interaction models for water in relation to protein hydration. In Pullman, B., editor, *Intermolecular Forces*, pages 331–342. D. Reidel Publishing Company, Dordrecht.

Berendsen, H., van der Spoel, D., and van Drunen, R. (1995). GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91(1-3):43–56.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., et al. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690.

Berglund, N. A., Piggot, T. J., Jefferies, D., et al. (2015). Interaction of the Antimicrobial Peptide Polymyxin B1 with Both Membranes of E. coli: A Molecular Dynamics Study. *PLOS Computational Biology*, 11(4):e1004180.

Berman, H. M., Westbrook, J., Feng, Z., et al. (2000). The protein data bank. Nucleic Acids Research, 28: 235-242. `http://www.rcsb.org/`.

Bernardi, R. C., Melo, M. C. R., and Schulten, K. (2015). Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1850(5):872–7.

Best, R. B., Clarke, J., and Karplus, M. (2005). What Contributions to Protein Sidechain Dynamics are Probed by NMR Experiments? A Molecular Dynamics Simulation Analysis. *Journal of Molecular Biology*, 349(1):185–203.

Bhadra, P., Yan, J., Li, J., et al. (2018). AmPEP: Sequence-based prediction of antimicrobial peptides using distribution patterns of amino acid properties and random forest. *Scientific Reports*, 8(1):1697.

Biasini, M., Mariani, V., Haas, J., et al. (2010). OpenStructure: a flexible software framework for computational structural biology. *Bioinformatics*, 26(20):2626–2628.

Billal, D. S., Feng, J., Leprohon, P., et al. (2011). Whole genome analysis of linezolid resistance in Streptococcus pneumoniae reveals resistance and compensatory mutations. *BMC Genomics*, 12(1):512.

Blair, J. M. A., Webber, M. A., Baylay, A. J., et al. (2015). Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*, 13(1):42–51.

Bland, J. M., De Lucca, A. J., Jacks, T. J., et al. (2001). All-D-cecropin B: Synthesis, conformation, lipopolysaccharide binding, and antibacterial activity. *Molecular and Cellular Biochemistry*, 218(1/2):105–111.

Bochicchio, D. and Pavan, G. M. (2017). From Cooperative Self-Assembly to Water-Soluble Supramolecular Polymers Using Coarse-Grained Simulations. *ACS Nano*, 11(1):1000–1011.

Böckmann, R. A., de Groot, B. L., Kakorin, S., et al. (2008). Kinetics, Statistics, and Energetics of Lipid Membrane Electroporation Studied by Molecular Dynamics Simulations. *Biophysical Journal*, 95(4):1837–1850.

Böckmann, R. A., Hac, A., Heimburg, T., et al. (2003). Effect of Sodium Chloride on a Lipid Bilayer. *Biophysical Journal*, 85(3):1647–1655.

Boisselier, E. and Astruc, D. (2009). Gold nanoparticles in nanomedicine: preparations, imaging, diagnostics, therapies and toxicity. *Chemical Society Reviews*, 38(6):1759.

Borle, F. and Seelig, J. (1983). Hydration of Escherichia coli lipids: Deuterium T1 relaxation time studies of phosphatidylglycerol, phosphatidylethanolamine and phosphatidylcholine. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 735(1):131–136.

Borle, F. and Seelig, J. (1985). Ca2+ binding to phosphatidylglycerol bilayers as studied by differential scanning calorimetry and 2H- and 31P-nuclear magnetic resonance. *Chemistry and Physics of Lipids*, 36(3):263–283.

Botan, A., Favela-Rosales, F., Fuchs, P. F. J., et al. (2015). Toward Atomistic Resolution Structure of Phosphatidylcholine Headgroup and Glycerol Backbone at Different Ambient Conditions. *The Journal of Physical Chemistry B*, 119(49):15075–15088.

Bransden, B. H. and Joachain, C. J. (2003). *Physics of atoms and molecules*. Prentice Hall.

Braun, A. R., Brandt, E. G., Edholm, O., et al. (2011). Determination of electron density profiles and area from simulations of undulating membranes. *Biophysical Journal*, 100(9):2112–20.

Breger, J. C., Muttenthaler, M., Delehanty, J. B., et al. (2017). Nanoparticle cellular uptake by dendritic wedge peptides: achieving single peptide facilitated delivery. *Nanoscale*, 9(29):10447–10464.

Brocos, P., Mendoza-Espinosa, P., Castillo, R., et al. (2012). Multiscale molecular dynamics simulations of micelles: coarse-grain for self-assembly and atomic resolution for finer details. *Soft Matter*, 8(34):9005.

Brogden, K. A. (2005). Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nature Reviews Microbiology*, 3(3):238–250.

Büldt, G., Gally, H. U., Seelig, A., et al. (1978). Neutron diffraction studies on selectively deuterated phospholipid bilayers. *Nature*, 271(5641):182–184.

Bunker, A., Magarkar, A., and Viitala, T. (2016). Rational design of liposomal drug delivery systems, a review: Combined experimental and computational studies of lipid membranes, liposomes and their PEGylation. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1858(10):2334–2352.

Burley, S. and Petsko, G. (1986). Amino-aromatic interactions in proteins. *FEBS Letters*, 203(2):139–143.

Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1):014101.

Camley, B. A., Lerner, M. G., Pastor, R. W., et al. (2015). Strong influence of periodic boundary conditions on lateral diffusion in lipid bilayer membranes. *The Journal of Chemical Physics*, 143(24):243113.

Carpenter, T. S., Parkin, J., and Khalid, S. (2016). The Free Energy of Small Solute Permeation through the *Escherichia coli* Outer Membrane Has a Distinctly Asymmetric Profile. *The Journal of Physical Chemistry Letters*, 7(17):3446–3451.

Castelletto, V., de Santis, E., Alkassem, H., et al. (2016). Structurally plastic peptide capsules for synthetic antimicrobial viruses. *Chemical Science*, 7(3):1707–1711.

Chan, D. I., Prenner, E. J., and Vogel, H. J. (2006). Tryptophan- and arginine-rich antimicrobial peptides: Structures and mechanisms of action. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1758(9):1184–1202.

Chan, Y.-H. M. and Boxer, S. G. (2007). Model membrane systems and their applications. *Current Opinion in Chemical Biology*, 11(6):581–7.

Chandrasekhar, I., Kastenholz, M., Lins, R. D., et al. (2003). A consistent potential energy parameter set for lipids: dipalmitoylphosphatidylcholine as a benchmark of the GROMOS96 45A3 force field. *European Biophysics Journal*, 32(1):67–77.

Chandrasekhar, I., Oostenbrink, C., and van Gunsteren, W. F. (2004). Simulating the Physiological Phase of Hydrated DPPC Bilayers: The Ester Moiety. *Soft Materials*, 2(1):27–45.

Chandrasekhar, I. and van Gunsteren, W. F. (2001). Sensitivity of molecular dynamics simulations of lipids to the size of the ester carbon. *Current Science*, 81(10):1325–1327.

Chandrasekhar, Wilfred F. van Gunst, I. and van Gunsteren, W. F. (2002). A comparison of the potential energy parameters of aliphatic alkanes: molecular dynamics simulations of triacylglycerols in the alpha phase. *European Biophysics Journal*, 31(2):89–101.

Chemical Computing Group ULC (2018). Molecular operating environment (moe), 2013.08. 1010 Sherbooke St. West, Suite 910, Montreal, QC, Canada, H3A 2R7.

Chen, C., Starr, C. G., Troendle, E. P., et al. (2019). Simulation-guided rational *de novo* design of a small pore-forming antimicrobial peptide. *Journal of the American Chemical Society*, 141(12):4839–4848.

Cherkasov, A., Hilpert, K., Jenssen, H., et al. (2009). Use of Artificial Intelligence in the Design of Small Peptide Antibiotics Effective against a Broad Spectrum of Highly Antibiotic-Resistant Superbugs. *ACS Chemical Biology*, 4(1):65–74.

Chiu, S.-W., Clark, M., Balaji, V., et al. (1995). Incorporation of surface tension into molecular dynamics simulation of an interface: a fluid phase lipid bilayer membrane. *Biophysical Journal*, 69(4):1230–45.

Chiu, S.-W., Pandit, S. A., Scott, H. L., et al. (2009). An Improved United Atom Force Field for Simulation of Mixed Lipid Bilayers. *The Journal of Physical Chemistry B*, 113(9):2748–2763.

Chowdhary, J., Harder, E., Lopes, P. E. M., et al. (2013). A Polarizable Force Field of Dipalmitoylphosphatidylcholine Based on the Classical Drude Model for Molecular Dynamics Simulations of Lipids. *The Journal of Physical Chemistry B*, 117(31):9142–9160.

Cipcigan, F., Carrieri, A. P., Pyzer-Knapp, E. O., et al. (2018). Accelerating molecular discovery through data and physical sciences: Applications to peptide-membrane interactions. *The Journal of Chemical Physics*, 148(24):241744.

City of Vienna Climate Protection Programme (2012). The Viennese Database for Disinfectants (WIDES Database). `https://www.wien.gv.at/english/environment/protection/oekokauf/disinfectants/`.

Clarke, R. J. (2001). The dipole potential of phospholipid membranes and methods for its detection. *Advances in Colloid and Interface Science*, 89-90:263–281.

Cochran, A. G., Skelton, N. J., and Starovasnik, M. A. (2001). Tryptophan zippers: Stable, monomeric β-hairpins. *Proceedings of the National Academy of Sciences*, 98(10):5578–5583.

Cocucci, E., Kim, J., Bai, Y., et al. (2017). Role of Passive Diffusion, Transporters, and Membrane Trafficking-Mediated Processes in Cellular Drug Transport. *Clinical Pharmacology & Therapeutics*, 101(1):121–129.

Collu, F., Spiga, E., Lorenz, C. D., et al. (2015). Assembly of Influenza Hemagglutinin Fusion Peptides in a Phospholipid Bilayer by Coarse-grained Computer Simulations. *Frontiers in Molecular Biosciences*, 2:66.

Cooper, G. M. (2000). *Transport of Small Molecules*. Sinauer Associates.

Darr, L., Machado, M. R., Brandner, A. F., et al. (2015). Sirah: A structurally unbiased coarse-grained force field for proteins with aqueous solvation and long-range electrostatics. *Journal of Chemical Theory and Computation*, 11(2):723–739. PMID: 26575407.

Davis, P. J., Fleming, B. D., Coolbear, K. P., et al. (1981). Gel to liquid-crystalline transition temperatures of water dispersions of two pairs of positional isomers of unsaturated mixed-acid phosphatidylcholines. *Biochemistry*, 20(12):3633–3636.

Daya, S. and Berns, K. I. (2008). Gene therapy using adeno-associated virus vectors. *Clinical Microbiology Reviews*, 21(4):583–93.

de Jong, D. H., Singh, G., Bennett, W. F. D., et al. (2013). Improved Parameters for the Martini Coarse-Grained Protein Force Field. *Journal of Chemical Theory and Computation*, 9(1):687–697.

de Vries, A. H., Mark, A. E., and Marrink, S.-J. (2004). Molecular Dynamics Simulation of the Spontaneous Formation of a Small DPPC Vesicle in Water in Atomistic Detail. *Journal of the American Chemical Society*, 126(14):4488–4489.

Delcour, A. H. (2009). Outer membrane permeability and antibiotic resistance. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1794(5):808–816.

Depan, D., Shah, J., and Misra, R. (2011). Controlled release of drug from folate-decorated and graphene mediated drug delivery system: Synthesis, loading efficiency, and drug release response. *Materials Science and Engineering: C*, 31(7):1305–1312.

Deslouches, B., Phadke, S. M., Lazarevic, V., et al. (2005). De Novo Generation of Cationic Antimicrobial Peptides: Influence of Length and Tryptophan Substitution on Antimicrobial Activity. *Antimicrobial Agents and Chemotherapy*, 49(1):316–322.

Dickson, C. J., Madej, B. D., Skjevik, Å. A., et al. (2014). Lipid14: The Amber Lipid Force Field. *Journal of Chemical Theory and Computation*, 10(2):865–879.

Ding, W., Palaiokostas, M., Wang, W., et al. (2015). Effects of Lipid Composition on Bilayer Membranes Quantified by All-Atom Molecular Dynamics. *The Journal of Physical Chemistry B*, 119(49):15263–15274.

Doktorova, M., Harries, D., and Khelashvili, G. (2017). Determination of bending rigidity and tilt modulus of lipid membranes from real-space fluctuation analysis of molecular dynamics simulations. *Physical Chemistry Chemical Physics*, 19(25):16806–16818.

Dolejska, M., Villa, L., Poirel, L., et al. (2013). Complete sequencing of an IncHI1 plasmid encoding the carbapenemase NDM-1, the ArmA 16S RNA methylase and a resistance-nodulation-cell division/multidrug efflux pump. *Journal of Antimicrobial Chemotherapy*, 68(1):34–39.

Douliez, J.-P., Ferrarini, A., and Dufourc, E.-J. (1998). On the relationship between C-C and C-D order parameters and its use for studying the conformation of lipid acyl chains in biomembranes. *The Journal of Chemical Physics*, 109(6):2513–2518.

Douliez, J. P., Léonard, A., and Dufourc, E. J. (1995). Restatement of order parameters in biomembranes: calculation of C-C bond order parameters from C-D quadrupolar splittings. *Biophysical Journal*, 68(5):1727–39.

Doux, J. P. F., Hall, B. A., and Killian, J. A. (2012). How lipid headgroups sense the membrane environment: an application of N NMR. *Biophysical Journal*, 103(6):1245–53.

Dror, R. O., Dirks, R. M., Grossman, J., et al. (2012). Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annual Review of Biophysics*, 41(1):429–452.

Durrett, R. (2010). *Probability: theory and examples*. Cambridge University Press.

Ebbensgaard, A., Mordhorst, H., Overgaard, M. T., et al. (2015). Comparative Evaluation of the Antimicrobial Activity of Different Antimicrobial Peptides against a Range of Pathogenic Bacteria. *PLOS ONE*, 10(12):e0144611.

Ebenhan, T., Gheysens, O., Kruger, H. G., et al. (2014). Antimicrobial peptides: their role as infection-selective tracers for molecular imaging. *BioMed Research International*, 2014:867381.

Editor BBC (2016). Artificial intelligence: Go master Lee Se-dol wins against AlphaGo program - BBC News. `https://www.bbc.co.uk/news/technology-35797102`.

Einstein, A. and Fürt, R. (1956). *Investigations on the theory of Brownian movement*. Dover Publications.

Entova, S., Billod, J.-M., Swiecicki, J.-M., et al. (2018). Insights into the key determinants of membrane protein topology enable the identification of new monotopic folds. *eLife*, 7:e40889.

Essmann, U., Perera, L., Berkowitz, M. L., et al. (1995). A smooth particle mesh Ewald method. *The Journal of Chemical Physics*, 103(19):8577–8593.

Fagan, R. P. and Fairweather, N. F. (2014). Biogenesis and functions of bacterial S-layers.

Fan, T., Yu, X., Shen, B., et al. (2017). Peptide Self-Assembled Nanostructures for Drug Delivery Applications. *Journal of Nanomaterials*, 2017:1–16.

Farnaud, S. and Evans, R. W. (2003). Lactoferrin - a multifunctional protein with antimicrobial properties. *Molecular Immunology*, 40(7):395–405.

Farrotti, A., Conflitti, P., Srivastava, S., et al. (2017). Molecular Dynamics Simulations of the Host Defense Peptide Temporin L and Its Q3K Derivative: An Atomic Level View from Aggregation in Water to Bilayer Perturbation. *Molecules*, 22(7):1235.

Fjell, C. D., Hiss, J. A., Hancock, R. E. W., et al. (2011). Designing antimicrobial peptides: form follows function. *Nature Reviews Drug Discovery*, 11(1):37.

Floyd, J. L., Smith, K. P., Kumar, S. H., et al. (2010). LmrS Is a Multidrug Efflux Pump of the Major Facilitator Superfamily from Staphylococcus aureus. *Antimicrobial Agents and Chemotherapy*, 54(12):5406–5412.

Frederix, P. W. J. M., Patmanidis, I., and Marrink, S.-J. (2018). Molecular simulations of self-assembling bio-inspired supramolecular systems and their connection to experiments. *Chemical Society Reviews*, 47(10):3470–3489.

Gahr, M., Speer, C. P., Damerau, B., et al. (1991). Influence of Lactoferrin on the Function of Human Polymorphonuclear Leukocytes and Monocytes. *Journal of Leukocyte Biology*, 49(5):427–433.

Gan, L., Chen, S., and Jensen, G. J. (2008). Molecular organization of Gram-negative peptidoglycan. *Proceedings of the National Academy of Sciences*, 105(48):18953–18957.

Gibbs, J. W. (2010). *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*. Cambridge Library Collection - Mathematics. Cambridge University Press.

Gifford, J. L., Hunter, H. N., and Vogel, H. J. (2005). Lactoferricin. *Cellular and Molecular Life Sciences*, 62(22):2588–2598.

Glukhov, E., Stark, M., Burrows, L. L., et al. (2005). Basis for selectivity of cationic antimicrobial peptides for bacterial versus mammalian membranes. *The Journal of Biological Chemistry*, 280(40):33960–7.

Gonçalves, M. B., Dreyer, J., Lupieri, P., et al. (2013). Structural prediction of a rhodamine-based biosensor and comparison with biophysical data. *Physical Chemistry Chemical Physics*, 15(6):2177–2183.

Google (2016). Google AI Blog. `https://ai.googleblog.com/2019/`.

Gordon, E., Mouz, N., Duée, E., et al. (2000). The crystal structure of the penicillin-binding protein 2x from Streptococcus pneumoniae and its acyl-enzyme form: implication in drug resistance. *Journal of Molecular Biology*, 299(2):477–485.

Gowers, R., Linke, M., Barnoud, J., et al. (2016). MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. In *Proceedings of the 15th Python in Science Conference*, pages 98–105.

Graham, J. A., Essex, J. W., and Khalid, S. (2017). PyCGTOOL: Automated Generation of Coarse-Grained Molecular Dynamics Models from Atomistic Trajectories. *Journal of Chemical Information and Modeling*, 57(4):650–656.

Grime, J. M. A., Dama, J. F., Ganser-Pornillos, B. K., et al. (2016). Coarse-grained simulation reveals key features of HIV-1 capsid self-assembly. *Nature Communications*, 7:11568.

Gudlur, S., Sukthankar, P., Gao, J., et al. (2012). Peptide Nanovesicles Formed by the Self-Assembly of Branched Amphiphilic Peptides. *PLOS ONE*, 7(9).

Guo, C., Luo, Y., Zhou, R., et al. (2012). Probing the Self-Assembly Mechanism of Diphenylalanine-Based Peptide Nanovesicles and Nanotubes. *ACS Nano*, 6(5):3907–3918.

Gurtovenko, A. A. and Vattulainen, I. (2007). Molecular Mechanism for Lipid Flip-Flops. *The Journal of Physical Chemistry B*, 111(48):13554–13559.

Gurtovenko, A. A. and Vattulainen, I. (2009). Calculation of the electrostatic potential of lipid bilayers from molecular dynamics simulations: Methodological issues. *The Journal of Chemical Physics*, 130(21):215107.

Hadden, J. A. and Perilla, J. R. (2018). All-atom virus simulations. *Current Opinion in Virology*, 31:82–91.

Hallock, K. J., Lee, D.-K., and Ramamoorthy, A. (2003). MSI-78, an analogue of the magainin antimicrobial peptides, disrupts lipid bilayer structure via positive curvature strain. *Biophysical Journal*, 84(5):3052–60.

Hancock, R. E. W. and Sahl, H.-G. (2006). Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nature Biotechnology*, 24(12):1551–1557.

Haria, N. R., Monticelli, L., Fraternali, F., et al. (2014). Plasticity and conformational equilibria of influenza fusion peptides in model lipid bilayers. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838(4):1169–1179.

Hart, K., Foloppe, N., Baker, C. M., et al. (2012). Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *Journal of Chemical Theory and Computation*, 8(1):348–362.

He, X., Song, M., Liang, H., et al. (2001). Self-assembly of the symmetric diblock copolymer in a confined state: Monte Carlo simulation. *The Journal of Chemical Physics*, 114(23):10510–10513.

Heo, L. and Feig, M. (2018). Experimental accuracy in protein structure refinement via molecular dynamics simulations. *Proceedings of the National Academy of Sciences*, 115(52):13276–13281.

Herce, H., Garcia, A., Litt, J., et al. (2009). Arginine-Rich Peptides Destabilize the Plasma Membrane, Consistent with a Pore Formation Translocation Mechanism of Cell-Penetrating Peptides. *Biophysical Journal*, 97(7):1917–1925.

Hess, B., Bekker, H., Berendsen, H. J. C., et al. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12):1463–1472.

Hilpert, K., Volkmer-Engert, R., Walter, T., et al. (2005). High-throughput generation of small antibacterial peptides with improved activity. *Nature Biotechnology*, 23(8):1008–1012.

Hodder, A. N., Crewther, P. E., and Anders, R. F. (2001). Specificity of the protective antibody response to apical membrane antigen 1. *Infection and Immunity*, 69(5):3286–94.

Hsu, P.-C., Samsudin, F., Shearer, J., et al. (2017). It Is Complicated: Curvature, Diffusion, and Lipid Sorting within the Two Membranes of Escherichia coli. *The Journal of Physical Chemistry Letters*, 8(22):5513–5518.

Huang, C. H., Lapides, J. R., and Levin, I. W. (1982). Phase-transition behavior of saturated, symmetric chain phospholipid bilayer dispersions determined by Raman spectroscopy: correlation between spectral and thermodynamic parameters. *Journal of the American Chemical Society*, 104(22):5926–5930.

Huang, J. and MacKerell, A. D. (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry*, 34(25):2135–2145.

Huang, J., Rauscher, S., Nawrocki, G., et al. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods*, 14(1):71–73.

Hughes, G. A. (2005). Nanostructure-mediated drug delivery. *Nanomedicine: Nanotechnology, Biology and Medicine*, 1(1):22–30.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14:33–38.

Hunter, C. A., Lawson, K. R., Perkins, J., et al. (2001). Aromatic interactions. *Journal of the Chemical Society, Perkin Transactions 2*, (5):651–669.

Hwang, P. M., Zhou, N., Shan, X., et al. (1998). Three-Dimensional Solution Structure of Lactoferricin B, an Antimicrobial Peptide Derived from Bovine Lactoferrin. *Biochemistry*, 37(12):4288–4298.

Institut Pasteur de Montevideo Group of BioMolecular Simulations (2019). Sirah force field for amino acids, solvent and nucleotides. `http://www.sirahff.com/2012/06/sirah-forcefield.html`.

Jahani, S., Shakiba, A., and Jahani, L. (2015). The Antimicrobial Effect of Lactoferrin on Gram-Negative and Gram-Positive Bacteria. *International Journal of Infection*, 2(3).

Jain, S. and Pillai, J. (2017). Bacterial membrane vesicles as novel nanosystems for drug delivery. *International Journal of Nanomedicine*, 12:6329–6341.

Jämbeck, J. P. M. and Lyubartsev, A. P. (2012). Derivation and Systematic Validation of a Refined All-Atom Force Field for Phosphatidylcholine Lipids. *The Journal of Physical Chemistry B*, 116(10):3164–3179.

Jarerattanachat, V., Karttunen, M., and Wong-ekkabut, J. (2013). Molecular Dynamics Study of Oxidized Lipid Bilayers in NaCl Solution. *The Journal of Physical Chemistry B*, 117(28):8490–8501.

Jean-Louis, S., Akare, S., Ali, M. A., et al. (2006). Deoxycholic acid induces intracellular signaling through membrane perturbations. *The Journal of Biological Chemistry*, 281(21):14948–60.

Jhong, J.-H., Chi, Y.-H., Li, W.-C., et al. (2019). dbAMP: an integrated resource for exploring antimicrobial peptides with functional activities and physicochemical properties on transcriptome and proteome data. *Nucleic Acids Research*, 47(D1):D285–D297.

Jia, Z., O'Mara, M. L., Zuegg, J., et al. (2011). The Effect of Environment on the Recognition and Binding of Vancomycin to Native and Resistant Forms of Lipid II. *Biophysical Journal*, 101(11):2684–2692.

Jiang, Q., Song, C., Nangreave, J., et al. (2012). DNA Origami as a Carrier for Circumvention of Drug Resistance. *Journal of the American Chemical Society*, 134(32):13396–13403.

Jiang, Z., Vasil, A. I., Gera, L., et al. (2011). Rational Design of α-Helical Antimicrobial Peptides to Target Gram-negative Pathogens, *Acinetobacter baumannii* and *Pseudomonas aeruginosa*: Utilization of Charge, 'Specificity Determinants', Total Hydrophobicity, Hydrophobe Type and Location as Design Parameters to Improve the Therapeutic Ratio. *Chemical Biology & Drug Design*, 77(4):225–240.

Jobin, M.-L., Vamparys, L., Deniau, R., et al. (2019). Biophysical Insight on the Membrane Insertion of an Arginine-Rich Cell-Penetrating Peptide. *International Journal of Molecular Sciences*, 20(18):4441.

Johner, N., Harries, D., and Khelashvili, G. (2016). Implementation of a methodology for determining elastic properties of lipid assemblies from molecular dynamics simulations. *BMC Bioinformatics*, 17(1):161.

Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996). Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society*, 118(45):11225–11236.

Jorgensen, W. L. and Tirado-Rives, J. (1988). The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6):1657–1666.

Joseleau-Petit, D., Liébart, J.-C., Ayala, J. A., et al. (2007). Unstable Escherichia coli L Forms Revisited: Growth Requires Peptidoglycan Synthesis. *Journal of Bacteriology*, 189(18):6512–6520.

Jusufi, A. (2013). Molecular simulations of self-assembly processes of amphiphiles in dilute solutions: the challenge for quantitative modelling. *Molecular Physics*, 111(21):3182–3192.

Kagan, B. L., Selsted, M. E., Ganz, T., et al. (1990). Antimicrobial defensin peptides form voltage-dependent ion-permeable channels in planar lipid bilayer membranes. *Proceedings of the National Academy of Sciences*, 87(1):210–214.

Kandasamy, S. K. and Larson, R. G. (2004). Binding and insertion of α-helical antimicrobial peptides in POPC bilayers studied by molecular dynamics simulations. *Chemistry and Physics of Lipids*, 132(1):113–132.

Kandt, C., Ash, W., and Peter Tieleman, D. (2007). Setting up and running molecular dynamics simulations of membrane proteins. *Methods*, 41(4):475–488.

Kepiro, I. E., Marzuoli, I., Hammond, K., et al. (2019). Engineering chirally blind protein pseudo-capsids into nanoprecise antibacterial persisters. Submitted.

Khalid, S., Berglund, N. A., Holdbrook, D. A., et al. (2015). The membranes of Gram-negative bacteria: progress in molecular modelling and simulation. *Biochemical Society Transactions*, 43(2):162–167.

Khalid, S., Piggot, T. J., and Samsudin, F. (2019). Atomistic and Coarse Grain Simulations of the Cell Envelope of Gram-Negative Bacteria: What Have We Learned? *Accounts of Chemical Research*, 52(1):180–188.

Khandelia, H. and Kaznessis, Y. N. (2005). Molecular dynamics simulations of helical antimicrobial peptides in SDS micelles: What do point mutations achieve? *Peptides*, 26(11):2037–2049.

Kim, J. (2017). Phage as a therapeutic agent. `https://medium.com/@thryve/phage-as-a-therapeutic-agent-ed4c466302e5`.

King, T. P., Wade, D., Coscia, M. R., et al. (1994). Structure-immunogenicity relationship of melittin, its transposed analogues, and D-melittin. *Journal of Immunology*, 153(3):1124–31.

Kirkpatrick, C. H., Green, I., Rich, R. R., et al. (1971). Inhibition of Growth of Candida albicans by Iron-Unsaturated Lactoferrin: Relation to Host-Defense Mechanisms in Chronic Mucocutaneous Candidiasis. *Journal of Infectious Diseases*, 124(6):539–544.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*, 220(4598):671–680.

Klauda, J. B., Venable, R. M., Freites, J. A., et al. (2010). Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *The Journal of Physical Chemistry B*, 114(23):7830–7843.

Kleandrova, V. V., Ruso, J. M., Speck-Planche, A., et al. (2016). Enabling the Discovery and Virtual Screening of Potent and Safe Antimicrobial Peptides. Simultaneous Prediction of Antibacterial Activity and Cytotoxicity. *ACS Combinatorial Science*, 18(8):490–498.

Kojima, S. and Nikaido, H. (2013). Permeation rates of penicillins indicate that Escherichia coli porins function principally as nonspecific channels. *Proceedings of the National Academy of Sciences*, 110(28):E2629–E2634.

Koziara, K., Stroet, M., Malde, A., et al. (2014). Testing and validation of the Automated Topology Builder (ATB) version 2.0: prediction of hydration free enthalpies. *Journal of Computer-Aided Molecular Design*, 28(3):221–233.

Krol, S. (2012). Challenges in drug delivery to the brain: Nature is against us. *Journal of Controlled Release*, 164(2):145–155.

Kučerka, N., Heberle, F., Pan, J., et al. (2015). Structural Significance of Lipid Diversity as Studied by Small Angle Neutron and X-ray Scattering. *Membranes*, 5(3):454–472.

Kučerka, N., Nieh, M.-P., and Katsaras, J. (2011). Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1808(11):2761–2771.

Kukol, A. (2009). Lipid Models for United-Atom Molecular Dynamics Simulations of Proteins. *Journal of Chemical Theory and Computation*, 5(3):615–626.

Kumariya, R., Sood, S. K., Rajput, Y. S., et al. (2015). Increased membrane surface positive charge and altered membrane fluidity leads to cationic antimicrobial peptide resistance in Enterococcus faecalis. *Biochimica et biophysica acta (BBA) - Biomembranes*, 1848(6):1367–75.

Ladokhin, A. S. and White, S. H. (2001). 'Detergent-like' permeabilization of anionic lipid vesicles by melittin. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1514(2):253–260.

Lai, Y. and Gallo, R. L. (2009). AMPed up immunity: how antimicrobial peptides have multiple roles in immune defense. *Trends in Immunology*, 30(3):131–141.

Lairion, F., Filler, R., and Disalvo, E. (2002). Reversed micelles as model systems to study interfacial properties of lipid bilayers. *Colloids and Surfaces B: Biointerfaces*, 25(4):369–371.

Lammers, T., Subr, V., Ulbrich, K., et al. (2009). Simultaneous delivery of doxorubicin and gemcitabine to tumors in vivo using prototypic polymeric drug carriers. *Biomaterials*, 30(20):3466–3475.

Lata, S., Mishra, N. K., and Raghava, G. P. (2010). AntiBP2: improved version of antibacterial peptide prediction. *BMC Bioinformatics*, 11(Suppl 1):S19.

Lauer, K. B., Borrow, R., and Blanchard, T. J. (2017). Multivalent and Multipathogen Viral Vector Vaccines. *Clinical and Vaccine Immunology*, 24(1):e00298–16.

Lavigne, J.-P., Sotto, A., Nicolas-Chanoine, M.-H., et al. (2013). An adaptive response of Enterobacter aerogenes to imipenem: regulation of porin balance in clinical isolates. *International Journal of Antimicrobial Agents*, 41(2):130–136.

Lee, E. H., Hsin, J., Sotomayor, M., et al. (2009). Discovery Through the Computational Microscope. *Structure*, 17(10):1295–1306.

Lee, H. and Pastor, R. W. (2011). Coarse-Grained Model for PEGylated Lipids: Effect of PEGylation on the Size and Shape of Self-Assembled Structures. *The Journal of Physical Chemistry B*, 115(24):7830–7837.

Lee, O.-S., Cho, V., and Schatz, G. C. (2012). Modeling the Self-Assembly of Peptide Amphiphiles into Fibers Using Coarse-Grained Molecular Dynamics. *Nano Letters*, 12(9):4907–4913.

Lehrer, R. I. (2004). Primate defensins. *Nature Reviews Microbiology*, 2(9):727–738.

Leontiadou, H., Mark, A. E., and Marrink, S.-J. (2004). Molecular dynamics simulations of hydrophilic pores in lipid bilayers. *Biophysical jJournal*, 86(4):2156–64.

Leontiadou, H., Mark, A. E., and Marrink, S.-J. (2006). Antimicrobial Peptides in Action. *Journal of the American Chemical Society*, 128(37):12156–12161.

Leontiadou, H., Mark, A. E., and Marrink, S.-J. (2007). Ion transport across transmembrane pores. *Biophysical Journal*, 92(12):4209–15.

Lewis, R. N. A. H., Mak, N., and McElhaney, R. N. (1987). A differential scanning calorimetric study of the thermotropic phase behavior of model membranes composed of phosphatidylcholines containing linear saturated fatty acyl chains. *Biochemistry*, 26(19):6118–6126.

Liang, J., Adamian, L., and Jackups, R. (2005). The membranewater interface region of membrane proteins: structural bias and the anti-snorkeling effect. *Trends in Biochemical Sciences*, 30(7):355–357.

Liechty, W. B., Kryscio, D. R., Slaughter, B. V., et al. (2010). Polymers for drug delivery systems. *Annual Review of Chemical and Biomolecular Engineering*, 1:149–73.

Lin, T.-Y. and Weibel, D. B. (2016). Organization and function of anionic phospholipids in bacteria. *Applied Microbiology and Biotechnology*, 100(10):4255–4267.

Lindblom, G. and Orädd, G. (2009). Lipid lateral diffusion and membrane heterogeneity. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1788(1):234–244.

Lipkin, R., Pino-Angeles, A., and Lazaridis, T. (2017). Transmembrane Pore Structures of β-Hairpin Antimicrobial Peptides by All-Atom Simulations. *The Journal of Physical Chemistry B*, 121(39):9126–9140.

Liu, S., Bao, J., Lao, X., et al. (2018). Novel 3D Structure Based Model for Activity Prediction and Design of Antimicrobial Peptides. *Scientific Reports*, 8(1):11189.

Liu, X., Liu, Y., and Liu, H. (2012). Theory and application of Monte Carlo method. In *Advances in Intelligent and Soft Computing*, volume 115 AISC, pages 841–848. Springer, Berlin, Heidelberg.

Livermore, D. (2008). Defining an extended-spectrum β-lactamase. *Clinical Microbiology and Infection*, 14:3–10.

Lobanovska, M. and Pilla, G. (2017). Penicillin's Discovery and Antibiotic Resistance: Lessons for the Future? *The Yale Journal of Biology and Medicine*, 90(1):135–145.

Lobo, F. P., Mota, B. E. F., Pena, S. D. J., et al. (2009). Virus-Host Coevolution: Common Patterns of Nucleotide Motif Usage in Flaviviridae and Their Hosts. *PLOS ONE*, 4(7):e6282.

Long, K. S., Poehlsgaard, J., Kehrenberg, C., et al. (2006). The Cfr rRNA Methyltransferase Confers Resistance to Phenicols, Lincosamides, Oxazolidinones, Pleuromutilins, and Streptogramin A Antibiotics. *Antimicrobial Agents and Chemotherapy*, 50(7):2500–2505.

Loose, C., Jensen, K., Rigoutsos, I., et al. (2006). A linguistic model for the rational design of antimicrobial peptides. *Nature*, 443(7113):867–869.

Lopes, D., Jakobtorweihen, S., Nunes, C., et al. (2017). Shedding light on the puzzle of drug-membrane interactions: Experimental techniques and molecular dynamics simulations. *Progress in Lipid Research*, 65:24–44.

Luo, T. and Robinson, D. N. (2015). Kinetic Monte Carlo simulations of the assembly of filamentous biomacromolecules by the dimer addition mechanism. *RSC Advances*, 5(6):3922–3929.

Ma, H., Cummins, D. D., Edelstein, N. B., et al. (2017). Modeling Diversity in Structures of Bacterial Outer Membrane Lipids. *Journal of Chemical Theory and Computation*, 13(2):811–824.

Ma, Y., Nolte, R. J., and Cornelissen, J. J. (2012). Virus-based nanocarriers for drug delivery. *Advanced Drug Delivery Reviews*, 64(9):811–825.

Mabrey, S. and Sturtevant, J. M. (1976). Investigation of phase transitions of lipids and lipid mixtures by sensitivity differential scanning calorimetry. *Proceedings of the National Academy of Sciences*, 73(11):3862–6.

Machado, M. R., González, H. C., and Pantano, S. (2017). MD Simulations of Viruslike Particles with Supra CG Solvation Affordable to Desktop Computers. *Journal of Chemical Theory and Computation*, 13(10):5106–5116.

Machado, M. R., Guisasola, E. E. B., Klein, F., et al. (2019). The SIRAH force field 2.0: Altius, Fortius, Citius. *Journal of Chemical Theory and Computation*, 15(4):2719–2733.

MacKerell, A. D., Bashford, D., Bellott, M., et al. (1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins . *The Journal of Physical Chemistry B*, 102(18):3586–3616.

Macpherson, J. A., Theisen, A., Masino, L., et al. (2019). Functional cross-talk between allosteric effects of activating and inhibiting ligands underlies PKM2 regulation. *eLife*, 8:e45068.

Mahlapuu, M., Håkansson, J., Ringstad, L., et al. (2016). Antimicrobial Peptides: An Emerging Category of Therapeutic Agents. *Frontiers in Cellular and Infection Microbiology*, 6:194.

Maier, J. A., Martinez, C., Kasavajhala, K., et al. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation*, 11(8):3696–3713.

Majumdar, B. B., Prytkova, V., Wong, E. K., et al. (2019). Role of Conformational Flexibility in Monte Carlo Simulations of Many-Protein Systems. *Journal of Chemical Theory and Computation*, 15(2):1399–1408.

Malanovic, N. (2016). Gram-positive bacterial cell envelopes: The impact on the activity of antimicrobial peptides. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1858(5):936–946.

Malay, A. D., Miyazaki, N., Biela, A., et al. (2019). An ultra-stable gold-coordinated protein cage displaying reversible assembly. *Nature*, 569(7756):438–442.

Malde, A., Zuo, L., Breeze, M., et al. (2011). An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *Journal of Chemical Theory and Computation*, 7(12):4026–4037.

Mälkiä, A., Murtomäki, L., Urtti, A., et al. (2004). Drug permeation in biomembranes: In vitro and in silico prediction and influence of physicochemical properties. *European Journal of Pharmaceutical Sciences*, 23(1):13–47.

Mantravadi, P., Kalesh, K., Dobson, R., et al. (2019). The Quest for Novel Antimicrobial Compounds: Emerging Trends in Research, Development, and Technologies. *Antibiotics*, 8(1):8.

Margreitter, C., Reif, M. M., and Oostenbrink, C. (2017). Update on phosphate and charged post-translationally modified amino acid parameters in the GROMOS force field. *Journal of Computational Chemistry*, 38(10):714–720.

Marinov, R. and Dufourc, E. J. (1996). Thermotropism and hydration properties of POPE and POPE-cholesterol systems as revealed by solid state2H and31P-NMR. *European Biophysics Journal*, 24(6):423–431.

Mark, P. and Nilsson, L. (2001). Structure and dynamics of the tip3p, spc, and spc/e water models at 298 k. *The Journal of Physical Chemistry A*, 105(43):9954–9960.

Marrink, S. J., de Vries, A. H., and Mark, A. E. (2004). Coarse grained model for semi-quantitative lipid simulations. *The Journal of Physical Chemistry B*, 108(2):750–760.

Marrink, S.-J., Lindahl, E., Edholm, O., et al. (2001). Simulation of the Spontaneous Aggregation of Phospholipids into Bilayers. *Journal of the American Chemical Society*, 123(35):8638–8639.

Marrink, S.-J. and Mark, A. E. (2003). The Mechanism of Vesicle Fusion as Revealed by Molecular Dynamics Simulations. *Journal of the American Chemical Society*, 125(37):11144–11145.

Marrink, S.-J., Risselada, H. J., Yefimov, S., et al. (2007). The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B*, 111(27):7812–7824.

Marrink, S.-J., Risselada, J., and Mark, A. E. (2005). Simulation of gel phase formation and melting in lipid bilayers using a coarse grained model. *Chemistry and Physics of Lipids*, 135(2):223–244.

Marrink, S.-J. and Tieleman, D. P. (2013). Perspective on the Martini model. *Chemical Society Reviews*, 42(16):6801–22.

Martínez, L., Andrade, R., Birgin, E. G., et al. (2009). PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry*, 30(13):2157–2164.

Marzuoli, I., Margreitter, C., and Fraternali, F. (2019). Lipid Head Group Parameterization for GROMOS 54A8: A Consistent Approach with Protein Force Field Description. *Journal of Chemical Theory and Computation*, 15(10):5175–5193.

Matsuzaki, K. (2009). Control of cell selectivity of antimicrobial peptides. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1788(8):1687–1692.

Matsuzaki, K., Murase, O., Fujii, N., et al. (1996). An Antimicrobial Peptide, Magainin 2, Induced Rapid Flip-Flop of Phospholipids Coupled with Pore Formation and Peptide Translocation. *Biochemistry*, 35(35):11361–11368.

Matsuzaki, K., Sugishita, K.-i., Harada, M., et al. (1997). Interactions of an antimicrobial peptide, magainin 2, with outer and inner membranes of Gram-negative bacteria. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1327(1):119–130.

McIntosh, T. J. and Simon, S. A. (1986). Area per molecule and distribution of water in fully hydrated dilauroylphosphatidylethanolamine bilayers. *Biochemistry*, 25(17):4948–4952.

Meißner, R. H., Schneider, J., Schiffels, P., et al. (2014). Computational Prediction of Circular Dichroism Spectra and Quantification of Helicity Loss upon Peptide Adsorption on Silica. *Langmuir*, 30(12):3487–3494.

Mettler Toledo (2018). Polymerization reactions. `https://www.mt.com/au/en/home/applications/L1_AutoChem_Applications/L2_ReactionAnalysis/L2_Polymerization.html#publications`.

Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., et al. (2011). MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry*, 32(10):2319–2327.

Migoń, D., Jaśkiewicz, M., Neubauer, D., et al. (2018). Alanine Scanning Studies of the Antimicrobial Peptide Aurein 1.2. *Probiotics and Antimicrobial Proteins*, pages 1–13.

Mills, M. and Andricioaei, I. (2008). An experimentally guided umbrella sampling protocol for biomolecules. *The Journal of Chemical Physics*, 129(11):114101.

Mingozzi, F. and High, K. A. (2011). Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nature Reviews Genetics*, 12(5):341–355.

Miyamoto, S. and Kollman, P. A. (1992). Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry*, 13(8):952–962.

Monticelli, L., Kandasamy, S. K., Periole, X., et al. (2008). The MARTINI Coarse-Grained Force Field: Extension to Proteins. *Journal of Chemical Theory and Computation*, 4(5):819–34.

Moskowitz, S. M., Ernst, R. K., and Miller, S. I. (2004). PmrAB, a two-component regulatory system of Pseudomonas aeruginosa that modulates resistance to cationic antimicrobial peptides and addition of aminoarabinose to lipid A. *Journal of Bacteriology*, 186(2):575–9.

Mulliken, R. S. (1955). Electronic Population Analysis on LCAOMO Molecular Wave Functions. I. *The Journal of Chemical Physics*, 23(10):1833–1840.

Naafs, M. A. B. (2018). The Antimicrobial Peptides: Ready for Clinical Trials? *Biomedical Journal of Scientific & Technical Research*, 7(4):001–005.

Nagle, J. F. (2017). Experimentally determined tilt and bending moduli of single-component lipid bilayers. *Chemistry and Physics of Lipids*, 205:18–24.

Nagle, J. F., Cognet, P., Dupuy, F. G., et al. (2019). Structure of gel phase DPPC determined by X-ray diffraction. *Chemistry and Physics of Lipids*, 218:168–177.

Nagle, J. F. and Tristram-Nagle, S. (2000). Structure of lipid bilayers. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, 1469(3):159–95.

Naldini, L. (2011). Ex vivo gene transfer and correction for cell-based therapies. *Nature Reviews Genetics*, 12(5):301–315.

Nelson, D. C., Garbe, J., and Collin, M. (2011). Cysteine proteinase SpeB from Streptococcus pyogenes - a potent modifier of immunologically important host and bacterial proteins. *Biological Chemistry*, 392(12):1077–88.

Nguyen, L. T., Haney, E. F., and Vogel, H. J. (2011). The expanding scope of antimicrobial peptide structures and their modes of action. *Trends in Biotechnology*, 29(9):464–472.

Nguyen, L. T., Schibli, D. J., and Vogel, H. J. (2005). Structural studies and model membrane interactions of two peptides derived from bovine lactoferricin. *Journal of Peptide Science*, 11(7):379–389.

Nicolas, J., Mura, S., Brambilla, D., et al. (2013). Design, functionalization strategies and biomedical applications of targeted biodegradable/biocompatible polymer-based nanocarriers for drug delivery. *Chemical Society Review*, 42(3):1147–1235.

Nikaido, E., Shirosaka, I., Yamaguchi, A., et al. (2011). Regulation of the AcrAB multidrug efflux pump in Salmonella enterica serovar Typhimurium in response to indole and paraquat. *Microbiology*, 157(3):648–655.

Nordmann, P., Poirel, L., Walsh, T. R., et al. (2011). The emerging NDM carbapenemases. *Trends in Microbiology*, 19(12):588–595.

Ogawa, W., Onishi, M., Ni, R., et al. (2012). Functional study of the novel multidrug efflux pump KexD from Klebsiella pneumoniae. *Gene*, 498(2):177–182.

Öjemalm, K., Higuchi, T., Lara, P., et al. (2016). Energetics of side-chain snorkeling in transmembrane helices probed by nonproteinogenic amino acids. *Proceedings of the National Academy of Sciences of the United States of America*, 113(38):10559–64.

O'neill, J. (2016). TACKLING DRUG-RESISTANT INFECTIONS GLOBALLY: FINAL REPORT AND RECOMMENDATIONS THE REVIEW ON ANTIMICROBIAL RESISTANCE. `https://amr-review.org/sites/default/files/160525{\_}Finalpaper{\_}withcover.pdf`.

Oostenbrink, C., Soares, T. A., van der Vegt, N. F. A., et al. (2005). Validation of the 53A6 GROMOS force field. *European Biophysics Journal*, 34(4):273–284.

Oostenbrink, C., Villa, A., Mark, A. E., et al. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, 25(13):1656–1676.

Orsi, M. (2018). Molecular simulation of self-assembly. *Self-assembling Biomaterials*, pages 305–318.

Pabst, G., Hodzic, A., Strancar, J., et al. (2007). Rigidification of neutral lipid bilayers in the presence of salts. *Biophysical Journal*, 93(8):2688–96.

Pan, J., Heberle, F. A., Tristram-Nagle, S., et al. (2012). Molecular structures of fluid phase phosphatidylglycerol bilayers as determined by small angle neutron and X-ray scattering. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1818(9):2135–2148.

Parrinello, M. and Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190.

Patra, M., Karttunen, M., Hyvönen, M., et al. (2003). Molecular Dynamics Simulations of Lipid Bilayers: Major Artifacts Due to Truncating Electrostatic Interactions. *Biophysical Journal*, 84(6):3636–3645.

Pattni, B. S., Chupin, V. V., and Torchilin, V. P. (2015). New Developments in Liposomal Drug Delivery. *Chemical Reviews*, 115(19):10938–10966.

Pattni, B. S. and Torchilin, V. P. (2015). Targeted drug delivery systems: Strategies and challenges. In Devarajan, P. V. and Jain, S., editors, *Targeted Drug Delivery: Concepts and Design*, pages 3–38. Springer International Publishing, Cham.

Perilla, J. R., Hadden, J. A., Goh, B. C., et al. (2016). All-Atom Molecular Dynamics of Virus Capsids as Drug Targets. *The Journal of Physical Chemistry Letters*, 7(10):1836–44.

Periole, X., Cavalli, M., Marrink, S.-J., et al. (2009). Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *Journal of Chemical Theory and Computation*, 5(9):2531–2543.

Peschel, A., Jack, R. W., Otto, M., et al. (2001). *Staphylococcus aureus* Resistance to Human Defensins and Evasion of Neutrophil Killing via the Novel Virulence Factor Mprf Is Based on Modification of Membrane Lipids with l-Lysine. *The Journal of Experimental Medicine*, 193(9):1067–1076.

Peschel, A. and Sahl, H.-G. (2006). The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nature Reviews Microbiology*, 4(7):529–536.

Petrache, H. I., Dodd, S. W., and Brown, M. F. (2000). Area per lipid and acyl length distributions in fluid phosphatidylcholines determined by (2)H NMR spectroscopy. *Biophysical Journal*, 79(6):3172–92.

Petrov, D., Margreitter, C., Grandits, M., et al. (2013). A Systematic Framework for Molecular Dynamics Simulations of Protein Post-Translational Modifications. *PLOS Computational Biology*, 9(7):e1003154.

Phillips, J. C., Braun, R., Wang, W., et al. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802.

Pickar, A. D. and Benz, R. (1978). Transport of oppositely charged lipophilic probe ions in lipid bilayer membranes having various structures. *The Journal of Membrane Biology*, 44(3-4):353–376.

Piggot, T. J., Allison, J. R., Sessions, R. B., et al. (2017). On the Calculation of Acyl Chain Order Parameters from Lipid Simulations. *Journal of Chemical Theory and Computation*, 13(11):5683–5696.

Piggot, T. J., Holdbrook, D. A., and Khalid, S. (2011). Electroporation of the E. coli and S. Aureus Membranes: Molecular Dynamics Simulations of Complex Bacterial Membranes. *The Journal of Physical Chemistry B*, 115(45):13381–13388.

Piggot, T. J., Piñeiro, Á., and Khalid, S. (2012). Molecular Dynamics Simulations of Phosphatidylcholine Membranes: A Comparative Force Field Study. *Journal of Chemical Theory and Computation*, 8(11):4593–4609.

Pino-Angeles, A., Leveritt, J. M., and Lazaridis, T. (2016). Pore Structure and Synergy in Antimicrobial Peptides of the Magainin Family. *PLOS Computational Biology*, 12(1):e1004570.

Pirtskhalava, M., Gabrielian, A., Cruz, P., et al. (2016). DBAASP v.2: an enhanced database of structure and antimicrobial/cytotoxic activity of natural and synthetic peptides. *Nucleic Acids Research*, 44(13):6503–6503.

Pluhackova, K., Kirsch, S. A., Han, J., et al. (2016). A Critical Comparison of Biomembrane Force Fields: Structure and Dynamics of Model DMPC, POPC, and POPE Bilayers. *The Journal of Physical Chemistry B*, 120(16):3888–3903.

Poger, D., Caron, B., and Mark, A. E. (2016). Validating lipid force fields against experimental data: Progress, challenges and perspectives. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1858(7):1556–1565.

Poger, D. and Mark, A. E. (2010). On the Validation of Molecular Dynamics Simulations of Saturated and cis -Monounsaturated Phosphatidylcholine Lipid Bilayers: A Comparison with Experiment. *Journal of Chemical Theory and Computation*, 6(1):325–336.

Poger, D. and Mark, A. E. (2012). Lipid Bilayers: The Effect of Force Field on Ordering and Dynamics. *Journal of Chemical Theory and Computation*, 8(11):4807–4817.

Poger, D., Van Gunsteren, W. F., and Mark, A. E. (2010). A new force field for simulating phosphatidylcholine bilayers. *Journal of Computational Chemistry*, 31(6):1117–1125.

Pokhrel, N. and Maibaum, L. (2018). Free Energy Calculations of Membrane Permeation: Challenges Due to Strong Headgroup-Solute Interactions. *Journal of Chemical Theory and Computation*, 14(3):1762–1771.

Poulou, A., Voulgari, E., Vrioni, G., et al. (2013). Outbreak Caused by an Ertapenem-Resistant, CTX-M-15-Producing Klebsiella pneumoniae Sequence Type 101 Clone Carrying an OmpK36 Porin Variant. *Journal of Clinical Microbiology*, 51(10):3176–3182.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raetz, C. R. H. and Whitfield, C. (2002). Lipopolysaccharide Endotoxins. *Annual Review of Biochemistry*, 71(1):635–700.

Rand, R. P., Fuller, N., Parsegian, V. A., et al. (1988). Variation in hydration forces between neutral phospholipid bilayers: evidence for hydration attraction. *Biochemistry*, 27(20):7711–7722.

Reif, M., Kallies, C., and Knecht, V. (2017). Effect of Sodium and Chloride Binding on a Lecithin Bilayer. A Molecular Dynamics Study. *Membranes*, 7(1):5.

Reif, M. M., Hünenberger, P. H., and Oostenbrink, C. (2012). New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *Journal of Chemical Theory and Computation*, 8(10):3705–3723.

Reif, M. M., Winger, M., and Oostenbrink, C. (2013). Testing of the GROMOS Force-Field Parameter Set 54A8: Structural Properties of Electrolyte Solutions, Lipid Bilayers, and Proteins. *Journal of Chemical Theory and Computation*, 9(2):1247–1264.

Reißer, S., Poger, D., Stroet, M., et al. (2017). Real Cost of Speed: The Effect of a Time-Saving Multiple-Time-Stepping Algorithm on the Accuracy of Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation*, 13(6):2367–2372.

Risselada, H. J. and Marrink, S.-J. (2008). The molecular face of lipid rafts in model membranes. *Proceedings of the National Academy of Sciences*, 105(45):17367–72.

Rossi, G., Manfrin, A., and Lutolf, M. P. (2018). Progress and potential in organoid research. *Nature Reviews Genetics*, 19(11):671–687.

Saar-Dover, R., Bitler, A., Nezer, R., et al. (2012). D-Alanylation of Lipoteichoic Acids Confers Resistance to Cationic Peptides in Group B Streptococcus by Increasing the Cell Wall Density. *PLOS Pathogens*, 8(9):e1002891.

Samsudin, F., Boags, A., Piggot, T. J., et al. (2017). Braun's Lipoprotein Facilitates OmpA Interaction with the Escherichia coli Cell Wall. *Biophysical Journal*, 113(7):1496–1504.

Sánchez, L., Calvo, M., and Brock, J. H. (1992). Biological role of lactoferrin. *Archives of Disease in Childhood*, 67(5):657–61.

Sandoval-Perez, A., Pluhackova, K., and Böckmann, R. A. (2017). Critical Comparison of Biomembrane Force Fields: ProteinLipid Interactions at the Membrane Interface. *Journal of Chemical Theory and Computation*, 13(5):2310–2321.

Santos, R., Ursu, O., Gaulton, A., et al. (2017). A comprehensive map of molecular drug targets. *Nature Reviews Drug Discovery*, 16(1):19–34.

Schamberger, J. and Clarke, R. J. (2002). Hydrophobic Ion Hydration and the Magnitude of the Dipole Potential. *Biophysical Journal*, 82(6):3081–3088.

Schibli, D. J., Hwang, P. M., and Vogel, H. J. (1999). The structure of the antimicrobial active center of lactoferricin B bound to sodium dodecyl sulfate micelles. *FEBS Letters*, 446(2-3):213–217.

Schmid, N., Eichenberger, A. P., Choutko, A., et al. (2011). Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *European Biophysics Journal*, 40(7):843–856.

Schoonen, L. and van Hest, J. C. M. (2014). Functionalization of protein-based nanocages for drug delivery applications. *Nanoscale*, 6(13):7124–7141.

Scott, D., Ghosh, A., Di, L., et al. (2017). Passive drug permeation through membranes and cellular distribution. *Pharmacological Research*, 117:94–102.

Scott, W., Hünenberger, P., Tironi, I., et al. (1999). The GROMOS Biomolecular Simulation Program Package. *The Journal of Physical Chemistry A*, 103(19):3596–3607.

Semchyschyn, D. J. and Macdonald, P. M. (2004). Conformational response of the phosphatidylcholine headgroup to bilayer surface charge: torsion angle constraints from dipolar and quadrupolar couplings in bicelles. *Magnetic Resonance in Chemistry*, 42(2):89–104.

Sengupta, D., Leontiadou, H., Mark, A. E., et al. (2008). Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1778(10):2308–2317.

Seo, M., Rauscher, S., Pomès, R., et al. (2012). Improving Internal Peptide Dynamics in the Coarse-Grained MARTINI Model: Toward Large-Scale Simulations of Amyloid- and Elastin-like Peptides. *Journal of Chemical Theory and Computation*, 8(5):1774–1785.

Sharpe, H. J., Stevens, T. J., and Munro, S. (2010). A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell*, 142(1):158–69.

Shau, H., Kim, A., and Golub, S. H. (1992). Modulation of natural killer and lymphokine-activated killer cell cytotoxicity by lactoferrin. *Journal of Leukocyte Biology*, 51(4):343–349.

Shi, Y. and Massagué, J. (2003). Mechanisms of TGF-β Signaling from Cell Membrane to the Nucleus. *Cell*, 113(6):685–700.

Shore, A. C., Deasy, E. C., Slickers, P., et al. (2011). Detection of Staphylococcal Cassette Chromosome *mec* Type XI Carrying Highly Divergent *mecA*, *mecI*, *mecR1*, *blaZ*, and *ccr* Genes in Human Clinical Isolates of Clonal Complex 130 Methicillin-Resistant *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, 55(8):3765–3773.

Sieprawska-Lupa, M., Mydel, P., Krawczyk, K., et al. (2004). Degradation of human antimicrobial peptide LL-37 by Staphylococcus aureus-derived proteinases. *Antimicrobial Agents and Chemotherapy*, 48(12):4673–9.

Silhavy, T. J., Kahne, D., and Walker, S. (2010). The Bacterial Cell Envelope. *Cold Spring Harbor Perspectives in Biology*, 2(5):a000414–a000414.

Silva, R. F., Araújo, D. R., Silva, E. R., et al. (2013). l-Diphenylalanine Microtubes As a Potential Drug-Delivery System: Characterization, Release Kinetics, and Cytotoxicity. *Langmuir*, 29(32):10205–10212.

Silva, T. F. D., Vila-Viçosa, D., Reis, P. B. P. S., et al. (2018). The Impact of Using Single Atomistic Long-Range Cutoff Schemes with the GROMOS 54A7 Force Field. *Journal of Chemical Theory and Computation*, 14(11):5823–5833.

Skjevik, Å. A., Madej, B. D., Dickson, C. J., et al. (2015). All-atom lipid bilayer self-assembly with the AMBER and CHARMM lipid force fields. *Chemical Communications*, 51(21):4402–4405.

Sohlenkamp, C. and Geiger, O. (2016). Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiology Reviews*, 40(1):133–159.

Song, C., de Groot, B. L., and Sansom, M. S. (2019). Lipid Bilayer Composition Influences the Activity of the Antimicrobial Peptide Dermcidin Channel. *Biophysical Journal*, 116(9):1658–1666.

Sørensen, S. J., Bailey, M., Hansen, L. H., et al. (2005). Studying plasmid horizontal transfer in situ: a critical review. *Nature Reviews Microbiology*, 3(9):700–710.

Spaar, A., Münster, C., and Salditt, T. (2004). Conformation of peptides in lipid membranes studied by x-ray grazing incidence scattering. *Biophysical Journal*, 87(1):396–407.

Spaeth, J. R., Kevrekidis, I. G., and Panagiotopoulos, A. Z. (2011). A comparison of implicit- and explicit-solvent simulations of self-assembly in block copolymer and solute systems. *The Journal of Chemical Physics*, 134(16):164902.

Spänig, S. and Heider, D. (2019). Encodings and models for antimicrobial peptide classification for multi-resistant pathogens. *BioData Mining*, 12(1):7.

Spector, A. A. and Yorek, M. A. (1985). Membrane lipid composition and cellular function. *Journal of Lipid Research*, 26(9):1015–35.

Starke-Peterkovic, T. and Clarke, R. J. (2009). Effect of headgroup on the dipole potential of phospholipid vesicles. *European Biophysics Journal*, 39(1):103–110.

Strebhardt, K. and Ullrich, A. (2008). Paul Ehrlich's magic bullet concept: 100 years of progress. *Nature Reviews Cancer*, 8(6):473–480.

Strøm, M. B., Haug, B. E., Rekdal, O., et al. (2002). Important structural features of 15-residue lactoferricin derivatives and methods for improvement of antimicrobial activity. *Biochemistry and Cell Biology*, 80(1):65–74.

Sun, D., Forsman, J., and Woodward, C. E. (2015). Atomistic Molecular Simulations Suggest a Kinetic Model for Membrane Translocation by Arginine-Rich Peptides. *The Journal of Physical Chemistry B*, 119(45):14413–14420.

Sun, R., Dama, J. F., Tan, J. S., et al. (2016). Transition-Tempered Metadynamics Is a Promising Tool for Studying the Permeation of Drug-like Molecules through Membranes. *Journal of Chemical Theory and Computation*, 12(10):5157–5169.

Sun, W.-J., Suter, R. M., Knewtson, M. A., et al. (1994). Order and disorder in fully hydrated unoriented bilayers of gel-phase dipalmitoylphosphatidylcholine. *Physical Review E*, 49(5):4665–4676.

Sunde, M., Serpell, L. C., Bartlam, M., et al. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology*, 273(3):729–739.

Swoboda, J. G., Campbell, J., Meredith, T. C., et al. (2009). Wall Teichoic Acid Function, Biosynthesis, and Inhibition. *ChemBioChem*, 11(1):35–45.

Takeuchi, K., Takahashi, H., Sugai, M., et al. (2004). Channel-forming membrane permeabilization by an antibacterial protein, sapecin: determination of membrane-buried and oligomerization surfaces by NMR. *The Journal of Biological Chemistry*, 279(6):4981–7.

Tang, M. and Hong, M. (2009). Structure and mechanism of β-hairpin antimicrobial peptides in lipid bilayers from solid-state NMR spectroscopy. *Molecular BioSystems*, 5(4):317.

Tang, M., Waring, A. J., and Hong, M. (2007). Phosphate-mediated arginine insertion into lipid membranes and pore formation by a cationic membrane peptide from solid-state nmr. *Journal of the American Chemical Society*, 129(37):11438–11446.

The Innovation Society (2019). Carbon Nanotubes can be Toxic to Aquatic Animals. `http://innovationsgesellschaft.ch/en/carbon-nanotubes-can-be-toxic-to-aquatic-animals/`.

The Johns Hopkins University (2017). Hopkins ABX Guide. `http://www.hopkins-abxguide.org/`.

Tieleman, D. and Marrink, S.-J. (2006). Lipids Out of Equilibrium: Energetics of Desorption and Pore Mediated Flip-Flop. *Journal of the American Chemical Society*, 128(38):12462–12467.

Tieleman, D. P. (2004). The molecular basis of electroporation. *BMC Biochemistry*, 5(1):10.

Tieleman, D. P. and Berendsen, H. J. C. (1998). Molecular dynamics simulations of a fully hydrated dipalmitoylphosphatidylcholine bilayer with different macroscopic boundary conditions and parameters. *The Journal of Chemical Physics*, 105(11):4871.

Tien, M. Z., Meyer, A. G., Sydykova, D. K., et al. (2013). Maximum Allowed Solvent Accessibilites of Residues in Proteins. *PLOS ONE*, 8(11):e80635.

Tironi, I. G., Sperb, R., Smith, P. E., et al. (1995). A generalized reaction field method for molecular dynamics simulations. *The Journal of Chemical Physics*, 102(13):5451–5459.

Tomita, M., Takase, M., Bellamy, W., et al. (1994). A review: the active peptide of lactoferrin. *Acta Paediatrica Japonica: Overseas edition*, 36(5):585–91.

Torres, M. D., Sothiselvam, S., Lu, T. K., et al. (2019). Peptide design principles for antimicrobial applications. *Journal of Molecular Biology*, 431(18):3547 – 3567. The molecular basis of antibiotic action and resistance.

Tsutsumi, A., Javkhlantugs, N., Kira, A., et al. (2012). Structure and orientation of bovine lactoferrampin in the mimetic bacterial membrane as revealed by solid-state NMR and molecular dynamics simulation. *Biophysical Journal*, 103(8):1735–43.

Ulmschneider, J. P. (2017). Charged Antimicrobial Peptides Can Translocate across Membranes without Forming Channel-like Pores. *Biophysical Journal*, 113(1):73–81.

Ulmschneider, M. B., Ulmschneider, J. P., Freites, J. A., et al. (2017). Transmembrane helices containing a charged arginine are thermodynamically stable. *European Biophysics Journal*, 46(7):627–637.

Ulrich, A. and Watts, A. (1994). Molecular response of the lipid headgroup to bilayer hydration monitored by 2H-NMR. *Biophysical Journal*, 66(5):1441–1449.

Ulrich, A. S., Volke, F., and Watts, A. (1990). The dependence of phospholipid headgroup mobility on hydration as studied by deuterium-NMR spin-lattice relaxation time measurements. *Chemistry and Physics of Lipids*, 55(1):61–66.

University of Calgary Biocomputing Group (2016). Image gallery. `https://www.ucalgary.ca/tieleman/gallery/image-gallery`.

Uppalapati, M., Lee, D. J., Mandal, K., et al. (2016). A Potent d -Protein Antagonist of VEGF-A is Nonimmunogenic, Metabolically Stable, and Longer-Circulating *in Vivo*. *ACS Chemical Biology*, 11(4):1058–1065.

Uppulury, K., Coppock, P. S., and Kindt, J. T. (2015). Molecular Simulation of the DPPE Lipid Bilayer Gel Phase: Coupling between Molecular Packing Order and Tail Tilt Angle. *The Journal of Physical Chemistry B*, 119:6.

van Gunsteren, W. F., Bakowies, D., Baron, R., et al. (2006). Biomolecular Modeling: Goals, Problems, Perspectives. *Angewandte Chemie International Edition*, 45(25):4064–4092.

van Gunsteren, W. F., Dolenc, J., and Mark, A. E. (2008). Molecular simulation as an aid to experimentalists. *Current Opinion in Structural Biology*, 18(2):149–153.

Van Lehn, R. C. and Alexander-Katz, A. (2014). Membrane-Embedded Nanoparticles Induce Lipid Rearrangements Similar to Those Exhibited by Biological Membrane Proteins. *The Journal of Physical Chemistry B*, 118(44):12586–12598.

Van Lehn, R. C., Ricci, M., Silva, P. H., et al. (2014). Lipid tail protrusions mediate the insertion of nanoparticles into model cell membranes. *Nature Communications*, 5(1):4482.

van Meer, G., Voelker, D. R., and Feigenson, G. W. (2008). Membrane lipids: where they are and how they behave. *Nature Reviews. Molecular cell biology*, 9(2):112–24.

Vargiu, A. V. and Nikaido, H. (2012). Multidrug binding properties of the AcrB efflux pump characterized by molecular dynamics simulations. *Proceedings of the National Academy of Sciences*, 109(50):20637–42.

Veltri, D., Kamath, U., and Shehu, A. (2018). Deep learning improves antimicrobial peptide recognition. *Bioinformatics*, 34(16):2740–2747.

Venable, R. M., Ingólfsson, H. I., Lerner, M. G., et al. (2017). Lipid and Peptide Diffusion in Bilayers: The Saffman-Delbrück Model and Periodic Boundary Conditions. *The Journal of Physical Chemistry B*, 121(15):3443–3457.

Vermeer, L. S., de Groot, B. L., Réat, V., et al. (2007). Acyl chain order parameter profiles in phospholipid bilayers: computation from molecular dynamics simulations and comparison with 2H NMR experiments. *European Biophysics Journal*, 36(8):919–931.

Vernier, P. T., Ziegler, M. J., Sun, Y., et al. (2006). Nanopore Formation and Phosphatidylserine Externalization in a Phospholipid Bilayer at High Transmembrane Potential. *Journal of the American Chemical Society*, 128(19):6288–6289.

Vivcharuk, V., Tomberli, B., Tolokh, I. S., et al. (2008). Prediction of binding free energy for adsorption of antimicrobial peptide lactoferricin B on a POPC membrane. *Physical Review E*, 77(3):031913.

Vogel, A. and Huster, D. (2017). Combining nmr spectroscopy and molecular dynamics simulation to investigate the structure and dynamics of membrane-associated proteins. In Chattopadhyay, A., editor, *Membrane Organization and Dynamics*, pages 311–350. Springer International Publishing, Cham.

Vögele, M., Köfinger, J., and Hummer, G. (2018). Hydrodynamics of Diffusion in Lipid Membrane Simulations. *Physical Review Letters*, 120(26):268104.

Wakabayashi, H., Abe, S., Okutomi, T., et al. (1996). Cooperative anti-Candida effects of lactoferrin or its peptides in combination with azole antifungal agents. *Microbiology and Immunology*, 40(11):821–5.

Wang, G., Li, X., Wang, Z., et al. (2016a). APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Research*, 44(D1):D1087–D1093.

Wang, H., Zhang, H., Liu, C., et al. (2012a). Coarse-grained molecular dynamics simulation of self-assembly of polyacrylamide and sodium dodecylsulfate in aqueous solution. *Journal of Colloid and Interface Science*, 386(1):205–211.

Wang, J., Chou, S., Xu, L., et al. (2015a). High specific selectivity and Membrane-Active Mechanism of the synthetic centrosymmetric α-helical peptides with Gly-Gly pairs. *Scientific Reports*, 5(1):15963.

Wang, J., Wolf, R. M., Caldwell, J. W., et al. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174.

Wang, L., Bose, P. S., and Sigworth, F. J. (2006a). Using cryo-EM to measure the dipole potential of a lipid membrane. *Proceedings of the National Academy of Sciences*, 103(49):18528–18533.

Wang, X., McGrath, S. C., Cotter, R. J., et al. (2006b). Expression cloning and periplasmic orientation of the Francisella novicida lipid A 4'-phosphatase LpxF. *The Journal of Biological Chemistry*, 281(14):9321–30.

Wang, Y., Chen, C. H., Hu, D., et al. (2016b). Spontaneous formation of structurally diverse membrane channel architectures from a single antimicrobial peptide. *Nature Communications*, 7:13535.

Wang, Y., Schlamadinger, D. E., Kim, J. E., et al. (2012b). Comparative molecular dynamics simulations of the antimicrobial peptide CM15 in model lipid bilayers. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1818(5):1402–1409.

Wang, Y., Zhao, T., Wei, D., et al. (2014). How reliable are molecular dynamics simulations of membrane active antimicrobial peptides? *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1838(9):2280–2288.

Wang, Z., Butner, J. D., Kerketta, R., et al. (2015b). Simulating cancer growth with multiscale agent-based modeling. *Seminars in Cancer Biology*, 30:70–78.

Wang, Z. Q., Lin, H. N., Li, S., et al. (1994). Calorimetric studies and molecular mechanics simulations of monounsaturated phosphatidylethanolamine bilayers. *The Journal of Biological Chemistry*, 269(38):23491–9.

Warshaviak, D. T., Muellner, M. J., and Chachisvilis, M. (2011). Effect of membrane tension on the electric field and dipole potential of lipid bilayer membrane. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1808(10):2608–2617.

Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., et al. (2015). Computational Lipidomics with *insane* : A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *Journal of Chemical Theory and Computation*, 11(5):2144–2155.

Wassenaar, T. A., Pluhackova, K., Böckmann, R. A., et al. (2014). Going backward: A flexible geometric approach to reverse transformation from coarse grained to atomistic models. *Journal of Chemical Theory and Computation*, 10(2):676–690. PMID: 26580045.

Wikipedia (2015). Liposome. `https://en.wikipedia.org/wiki/Liposome`.

Wilson, J. R., Clark, R. B., Banderali, U., et al. (2011). Measurement of the membrane potential in small cells using patch clamp methods. *Channels*, 5(6):530–537.

Wipf, P., Xiao, J., and Stephenson, C. R. J. (2009). Peptide-Like Molecules (PLMs): A Journey from Peptide Bond Isosteres to Gramicidin S Mimetics and Mitochondrial Targeting Agents. *CHIMIA International Journal for Chemistry*, 63(11):764–775.

Wishart, D. S., Feunang, Y. D., Guo, A. C., et al. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Research*, 46(D1):D1074–D1082.

Woodford, N. and Johnson, A. P. (2013). Global spread of antibiotic resistance: the example of New Delhi metallo-β-lactamase (NDM)-mediated carbapenem resistance. *Journal of Medical Microbiology*, 62(4):499–513.

Wright, G. (2005). Bacterial resistance to antibiotics: Enzymatic degradation and modification. *Advanced Drug Delivery Reviews*, 57(10):1451–1470.

Wu, D. and Yang, X. (2012). Coarse-Grained Molecular Simulation of Self-Assembly for Nonionic Surfactants on Graphene Nanostructures. *The Journal of Physical Chemistry B*, 116(39):12048–12056.

Wu, W., Hsiao, S. C., Carrico, Z. M., et al. (2009). Genome-free viral capsids as multivalent carriers for taxol delivery. *Angewandte Chemie (International ed. in English)*, 48(50):9493–7.

Yamaguchi, S., Huster, D., Waring, A., et al. (2001). Orientation and Dynamics of an Antimicrobial Peptide in the Lipid Bilayer by Solid-State NMR Spectroscopy. *Biophysical Journal*, 81(4):2203–2214.

Yang, L., Harroun, T. A., Weiss, T. M., et al. (2001). Barrel-stave model or toroidal model? A case study on melittin pores. *Biophysical Journal*, 81(3):1475–85.

Yeaman, M. R. and Yount, N. Y. (2003). Mechanisms of Antimicrobial Peptide Action and Resistance. *Pharmacological Reviews*, 55(1):27–55.

Yeates, T. O. (2019). Protein assembles into Archimedean geometry. *Nature*, 569(7756):340–342.

Yesylevskyy, S. O., Schäfer, L. V., Sengupta, D., et al. (2010). Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *PLOS Computational Biology*, 6(6):e1000810.

Yingchoncharoen, P., Kalinowski, D. S., and Richardson, D. R. (2016). Lipid-Based Drug Delivery Systems in Cancer Therapy: What Is Available and What Is Yet to Come. *Pharmacological Reviews*, 68(3):701–87.

Yonezawa, A., Kuwahara, J., Fujii, N., et al. (1992). Binding of tachyplesin I to DNA revealed by footprinting analysis: significant contribution of secondary structure to DNA binding and implication for biological action. *Biochemistry*, 31(11):2998–3004.

Zasloff, M. (2002). Antimicrobial peptides of multicellular organisms. *Nature*, 415(6870):389–395.

Zhang, L., Rozek, A., and Hancock, R. E. (2001). Interaction of cationic antimicrobial peptides with model membranes. *The Journal of Biological Chemistry*, 276(38):35714–22.

Zhang, Q., Jiang, Q., Li, N., et al. (2014). DNA Origami as an *In Vivo* Drug Delivery Vehicle for Cancer Therapy. *ACS Nano*, 8(7):6633–6643.

Zhao, L., Cao, Z., Bian, Y., et al. (2018a). Molecular Dynamics Simulations of Human Antimicrobial Peptide LL-37 in Model POPC and POPG Lipid Bilayers. *International Journal of Molecular Sciences*, 19(4):1186.

Zhao, R., Babani, S., Gao, F., et al. (2000). The mechanism of transport of the multitargeted antifolate (MTA) and its cross-resistance pattern in cells with markedly impaired transport of methotrexate. *Clinical Cancer Research: an official journal of the American Association for Cancer Research*, 6(9):3687–95.

Zhao, Y., Yang, W., Wang, D., et al. (2018b). Controlling the Diameters of Nanotubes Self-Assembled from Designed Peptide Bolaphiles. *Small*, 14(12):1703216.